

1/ 235

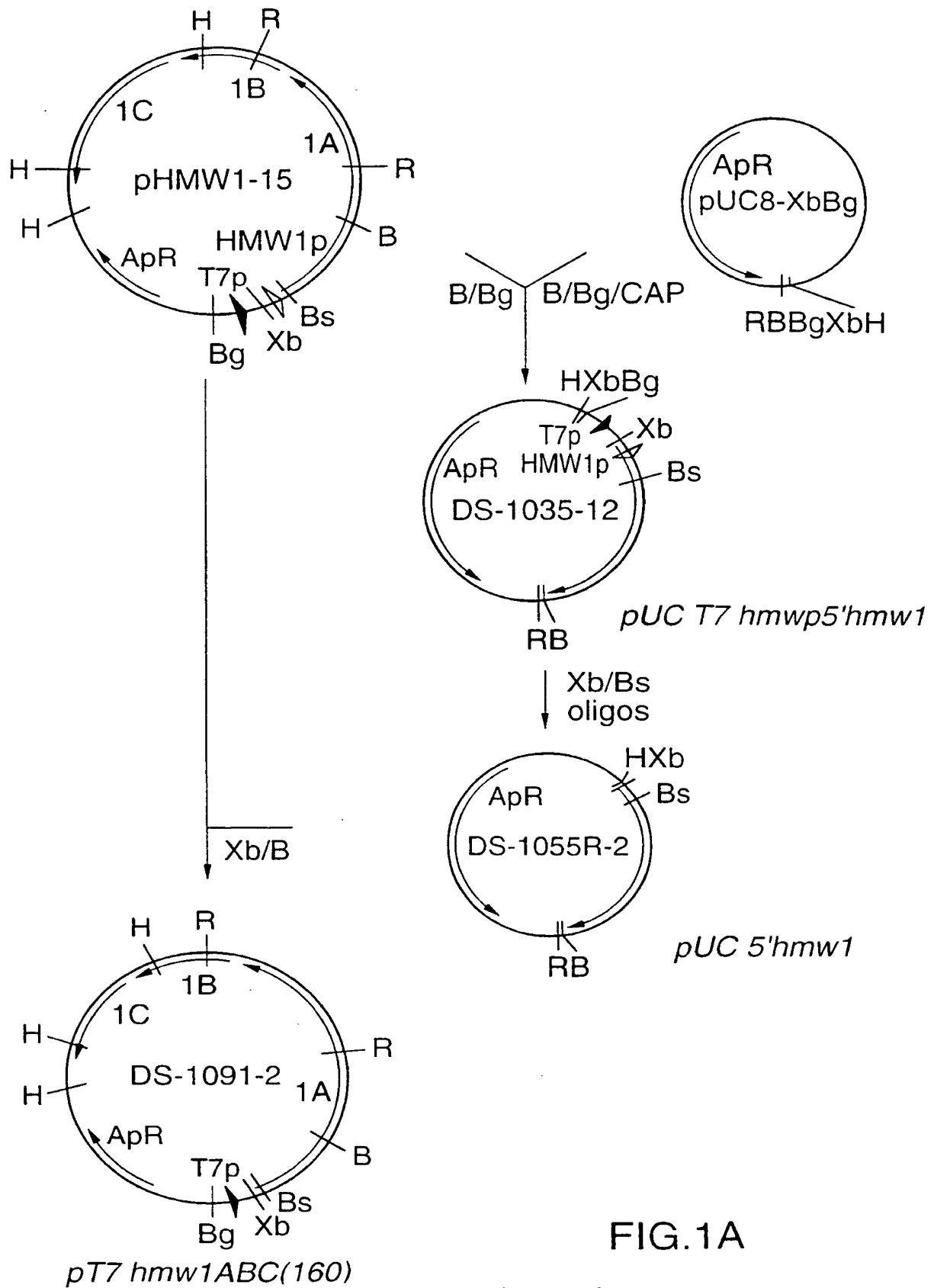
Construction of DS-1091-2, a *hmw1ABC* (160) plasmid.

FIG.1A

2/235

FIG.1B

Oligonucleotides to re-create the N-terminus of the full-length HMW1A protein in plasmid DS-1091-2 or the N-terminus of the full-length HMW2A protein in plasmid DS-1094-2.

<u>SEQ ID NO</u>	
Xba I	↓
↓	↓
CTAGAAATAATTTGTGTTTAACTTTAAGAGGAGATATACATATGAACAAGATATATCGTCTCAAAATTCAGCAAACGCCCTGAATGCT	1
TTTATTAAACAAATTGAAATTTCCTCTATATGTATATCTATATAGCAGAGTTTAAAGTCGTTTCGGGACTTAC	2
	3

3/ 235

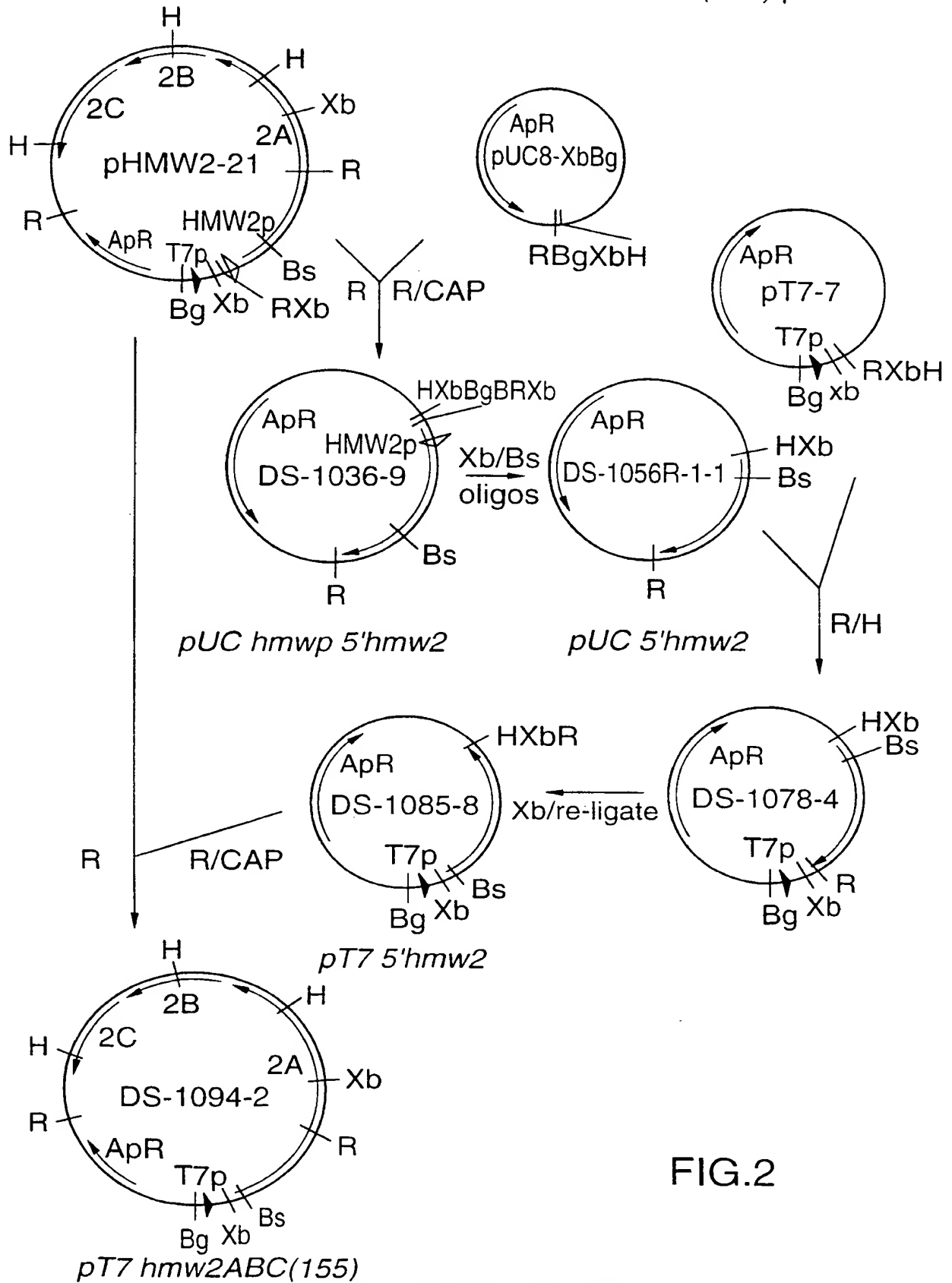
Construction of DS-1094-2, a T7 *hmw2ABC* (155) plasmid.

FIG.2

4/235

Construction of DS-1046-1-1, a T7 *hmw1ABC* (125) plasmid.

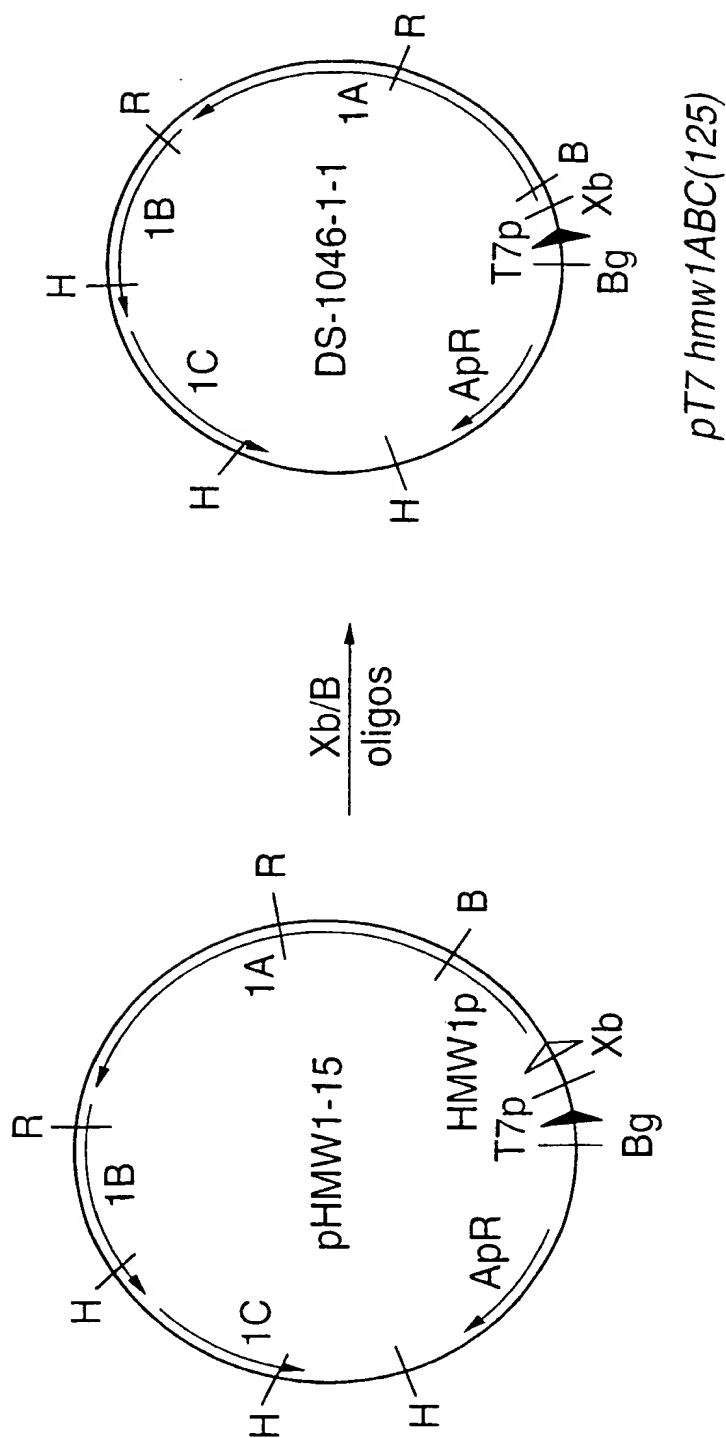


FIG.3A

FIG. 3B

Oligonucleotides to re-create the N-terminus of the mature HMMA protein in plasmid DS-1046-1-1.

SEO ID NO

Xba I

$$M \quad P \quad D \quad N \quad V \quad \dots$$

CTAGAAATAATTTTGTGTTAACTTTAAGAGGAGATATACATATGCGCGATAATGTAT...

TTTTATTAAACAATTGAAATTCCTCTATATGTATACGGCTATTACATA...

...S I N A E T A G R S N T S E D D E Y T B a n H I 4

→

...CTATTATTCAGAAACAGCAGGACGCAGCAATACTTTCAGAAAGACGATGAATACACGG

...GATAATTACGCTTTTGTCGTCCTGGTCGTTATGAAGCTTTCGCTACTTATGTGCCCCTAG

6/235

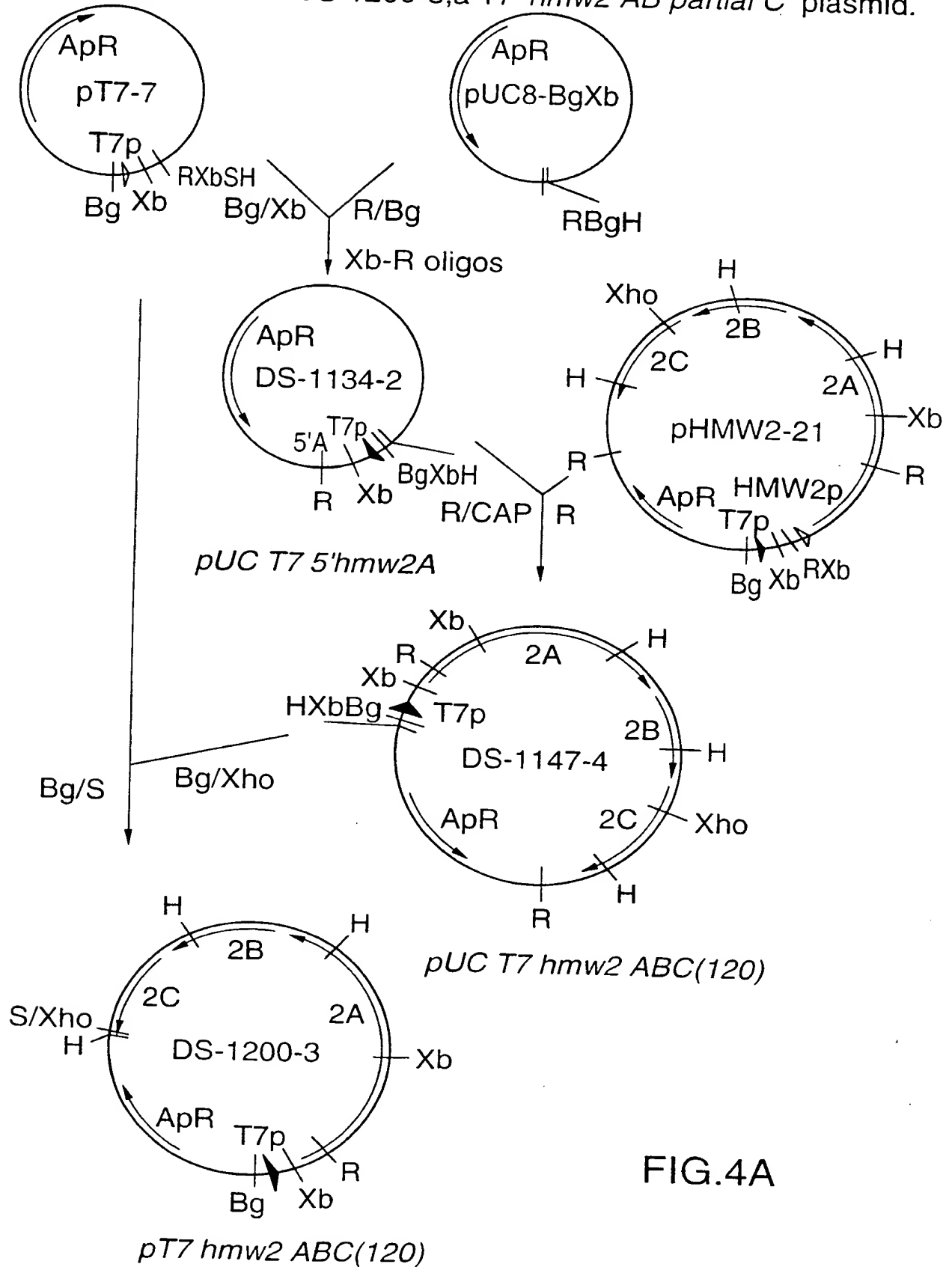
Construction of DS-1200-3, a T7 *hmw2* ABC partial C plasmid.

FIG.4A

7 / 235

SEQ ID NO

Oligonucleotides used to re-create the N-terminus of the HMW2A protein in plasmid DS-1200-3.

xIba I

M P D D ...

CTAGAAATAATTGTCTTAACCTTAAGAAGGAGATATACATATGCCCTGATGATG...

TTTTATTAAACAAATTGAAATTCCTCTATATGTATACGGACTACTAC...

...V T I E A E D P L R N N T G I N D E C O R I

7

...TAACAATTGAAGCCGAAGACCCCCCTTCGCCAATAATACCGGTATAAATGATG

...ATTGTTAACTTCGGCTTCGCGGAAGCGTTATTATGCCATATTACTATTAA

9

8/235

Construction of DS-1122-2, a T7 hmw1A partial B plasmid.

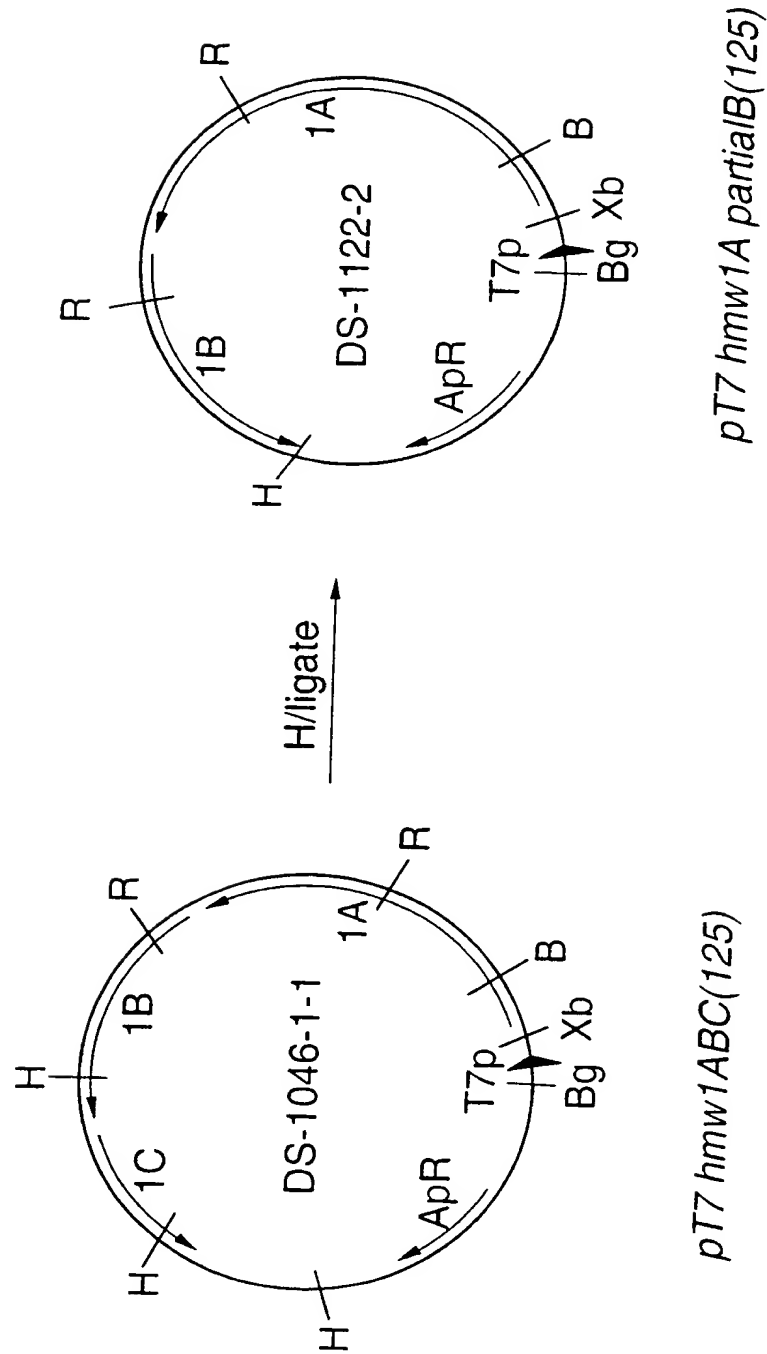


FIG.5

9/235

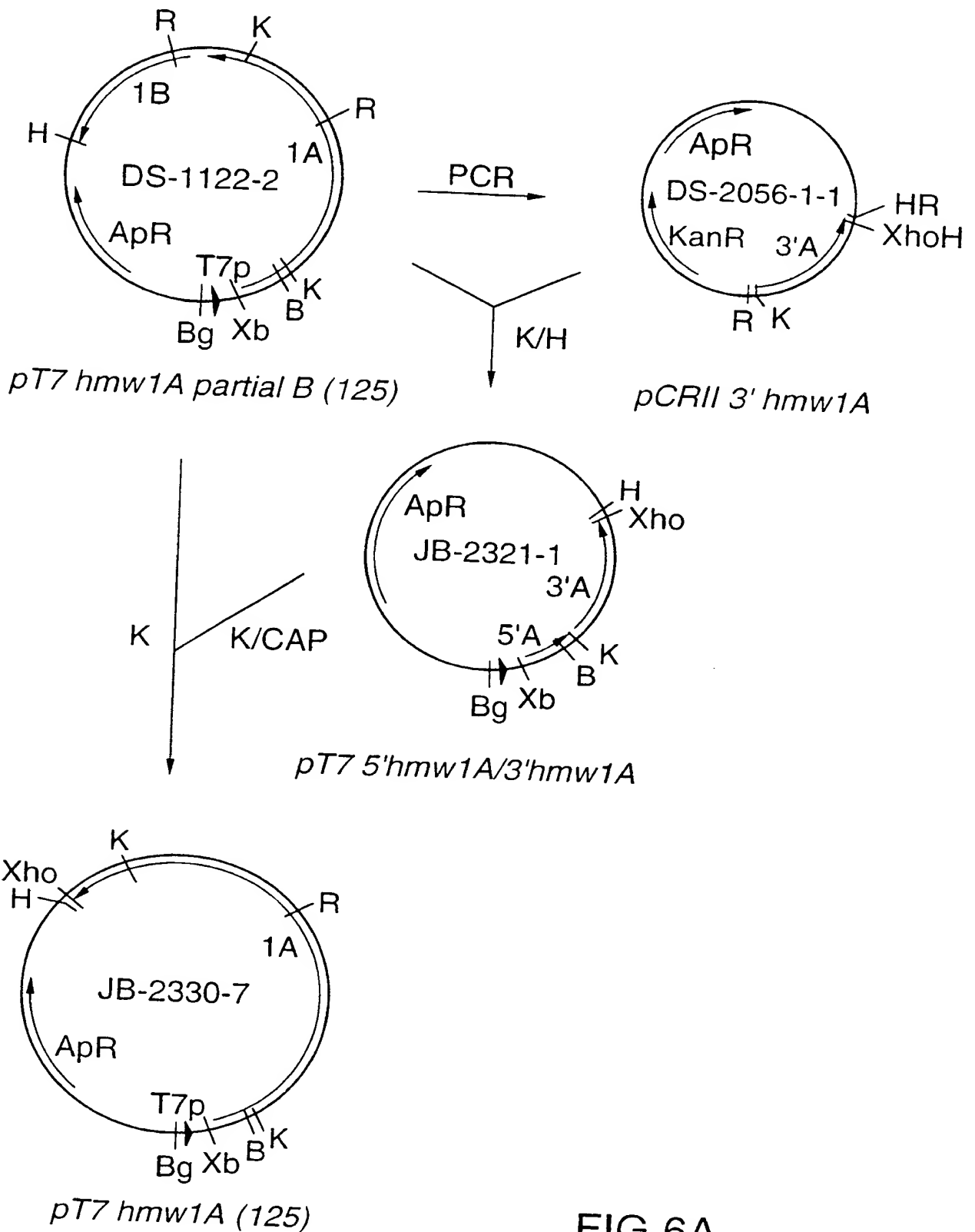
Construction of JB-2330-7, a *T7 hmw1A* (125) plasmid.

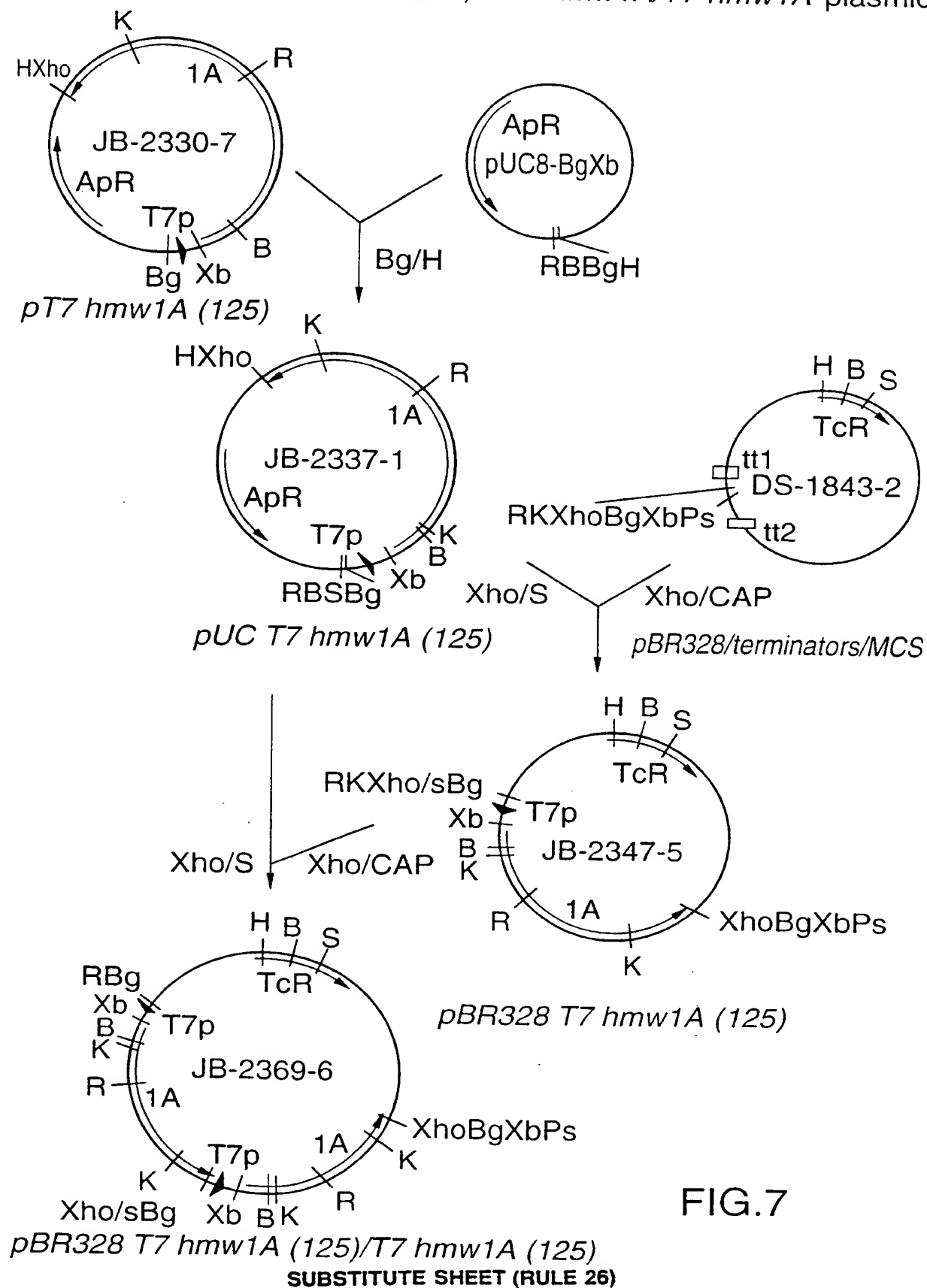
FIG.6A

FIG.6B

Oligonucleotides used to PCR amplify the 3'-end of *hmm1A*.

sense		SEO ID NO		10/235
	<div> <div> Kpn I ↓ </div> <div> I S G T L V I N A K D A E 5' C=ACCAGCGGTACCTTGGTTATTAAACGCAAAAGACGCTGAG 3' </div> </div>	10 11	5358.JS	
antisense	<div> <div> V N I A D N G R * CGGTTAATATCGCTGATAACGGCGGTAG 3' CCCAATTATAGCGGACTATTGCCCGCCCATCGAGCTCTTCGAACCGG 5' </div> <div> ↑ Xho I Hind III </div> </div>	12 13 14	5359.JS	

11/235

Construction of JB-2369-6, a *T7 hmw1A*/*T7 hmw1A* plasmid.

12/235

Construction of DS-2084-1 and DS-2084-3, plasmids with one or two T7 hmw2A genes

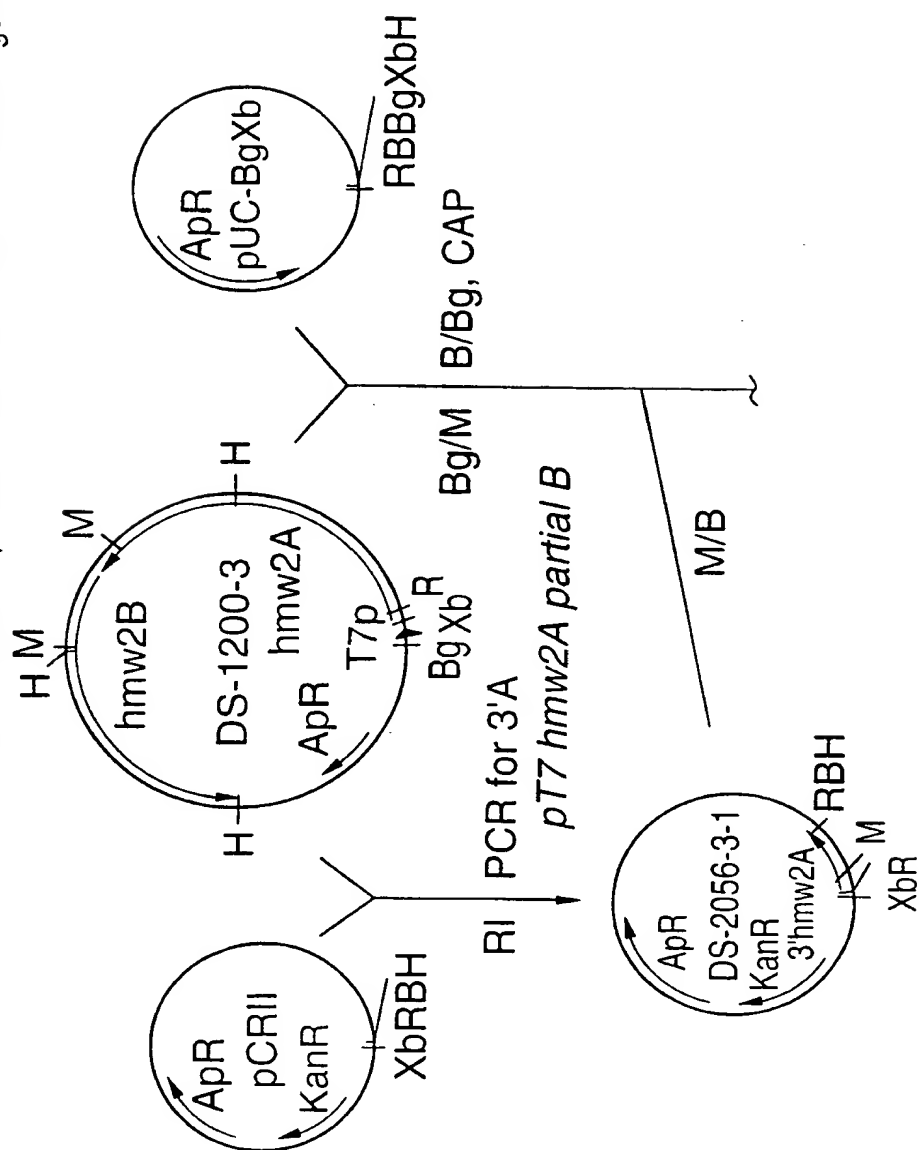


FIG.8A

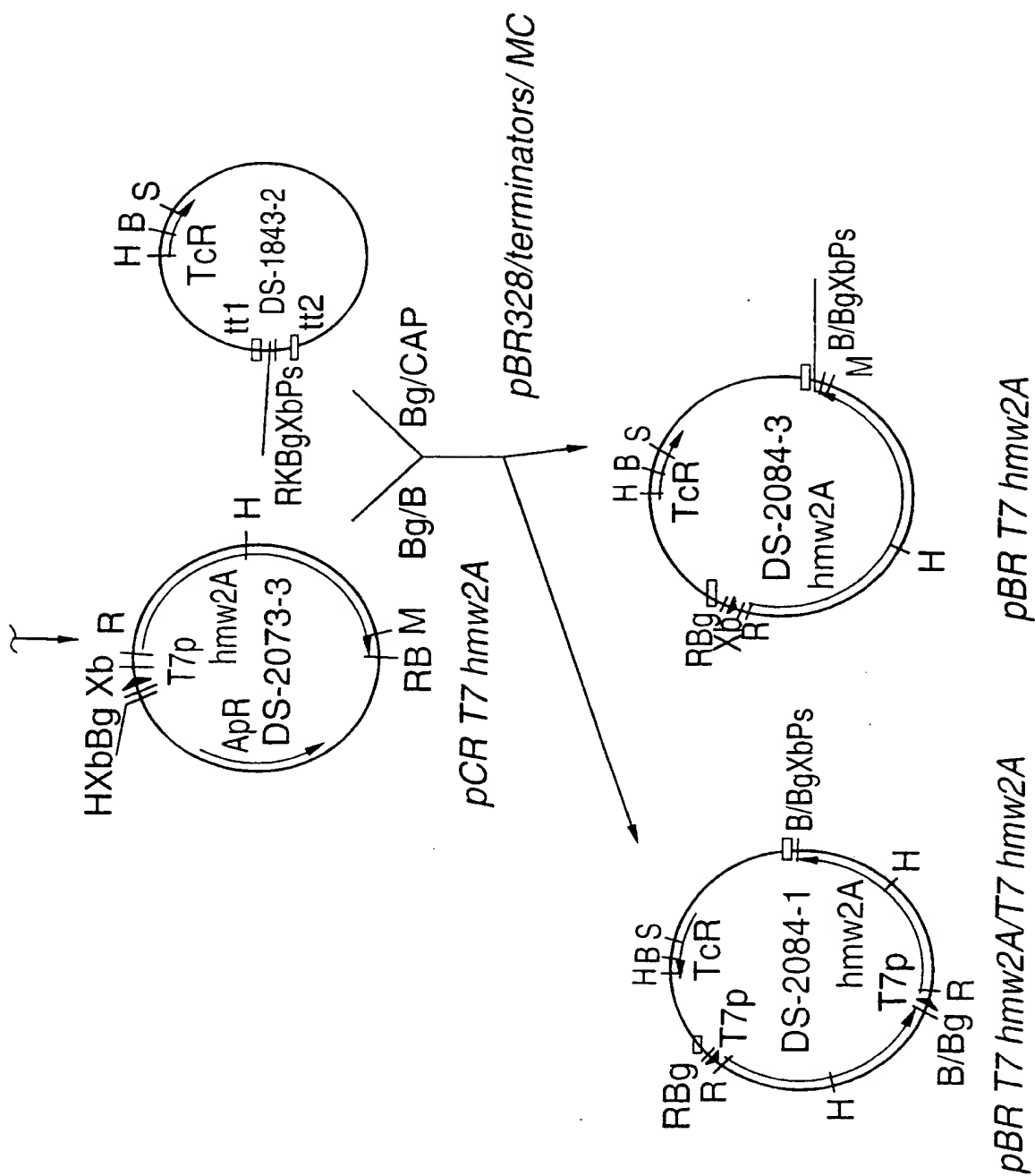


FIG.8A'

FIG. 8B

Oligonucleotides used to PCR amplify the 3'-end of *hmm2A*.

sense

ECOR I

5' CCGGAATTCGAAACGGCTCTTGAAAAAGTAAAAG K R V L E K V K

5360.DC

15 16

antisense

T N V A D D G Q P *

TACCAATGTTGCTGACGATGGACAGCCGTAG

3' ATGTTACAAACGACTGCTACTTACCTGTCTGGCATCTAGGGC 5' 5361.DC

5

5361.DC

17 18 19

BamH I

15/235

Construction of BK-86-1-1, a *T7 hmw1A/T7 hmw1ABC/KanR* plasmid.

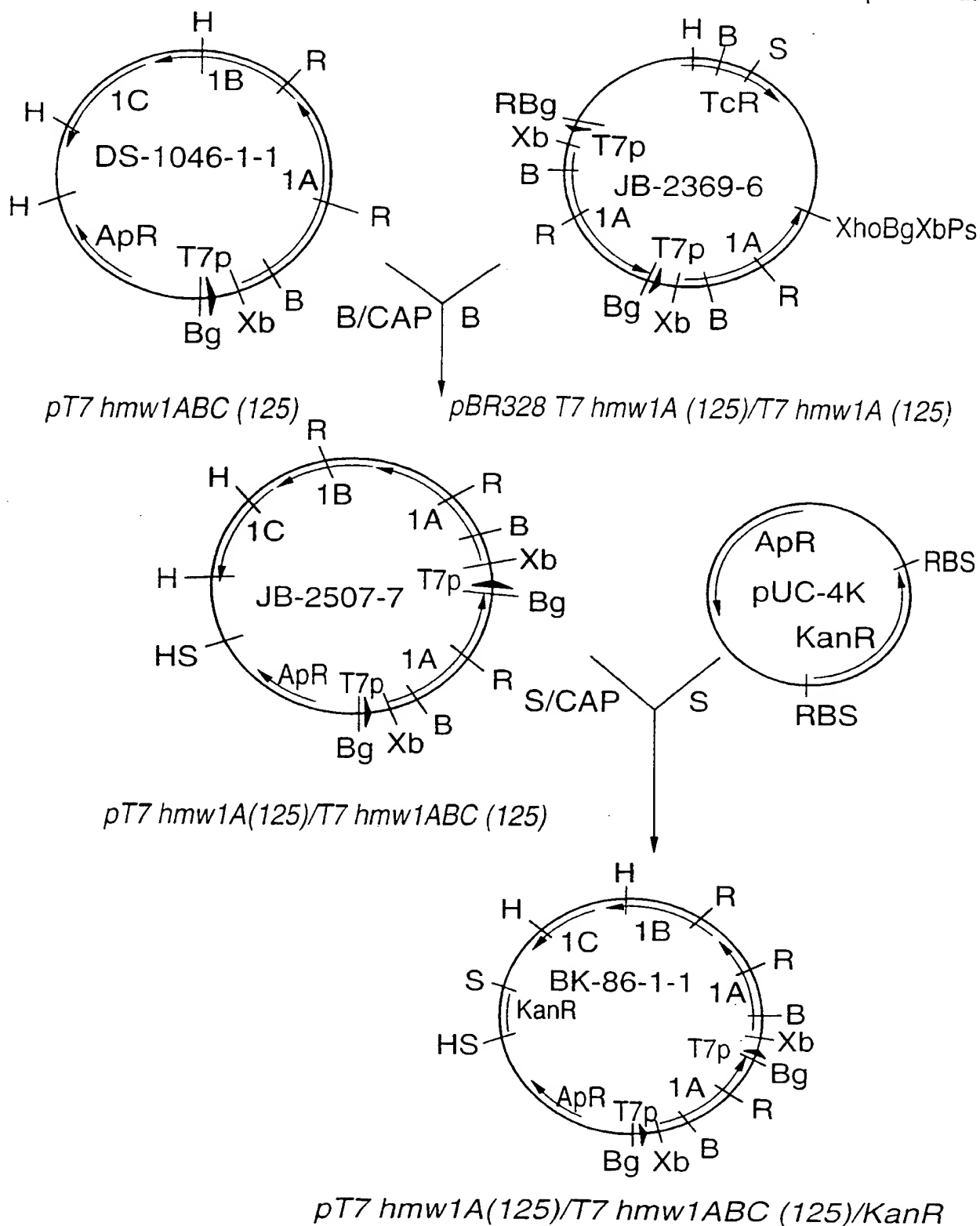


FIG.9

16/235

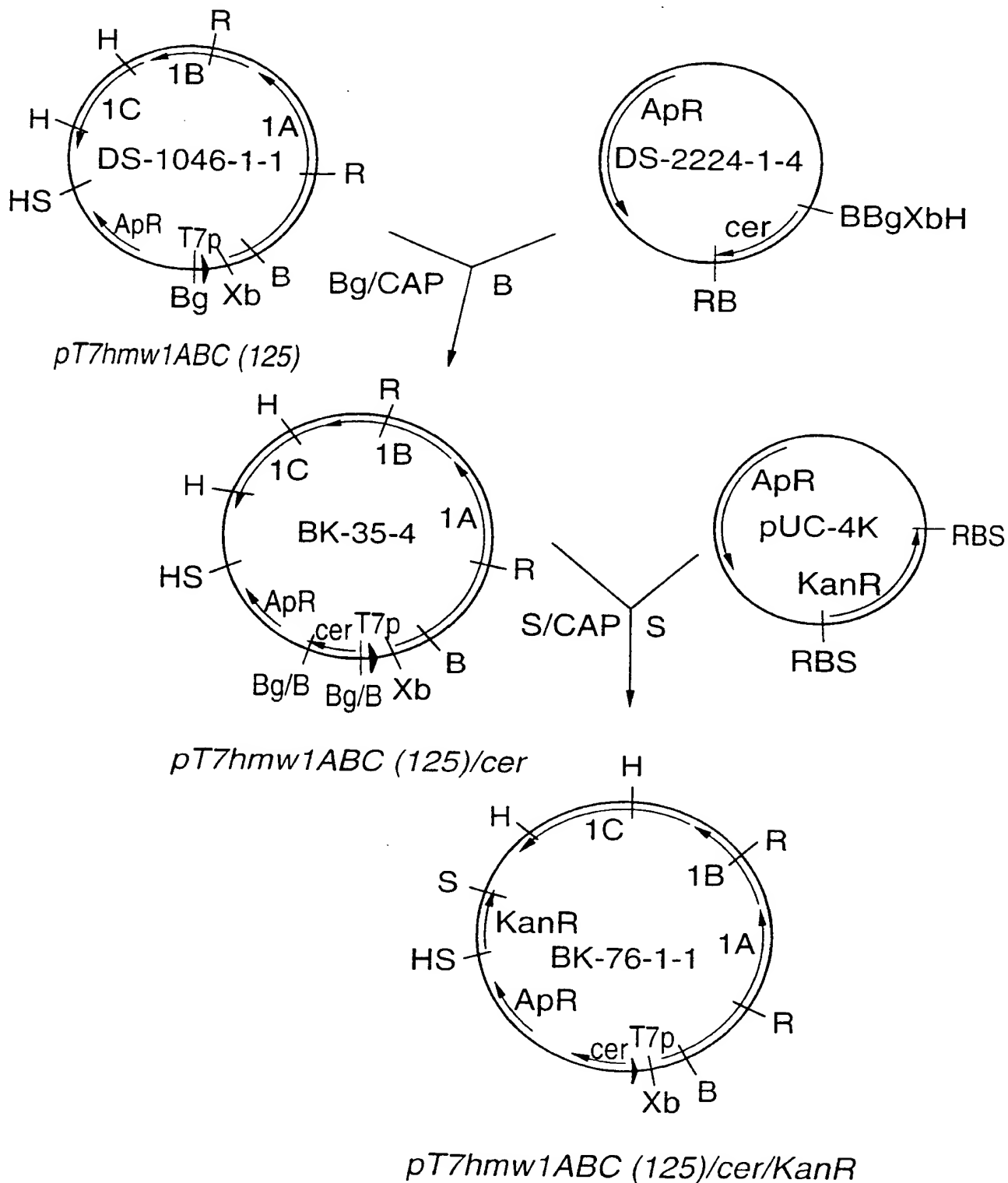
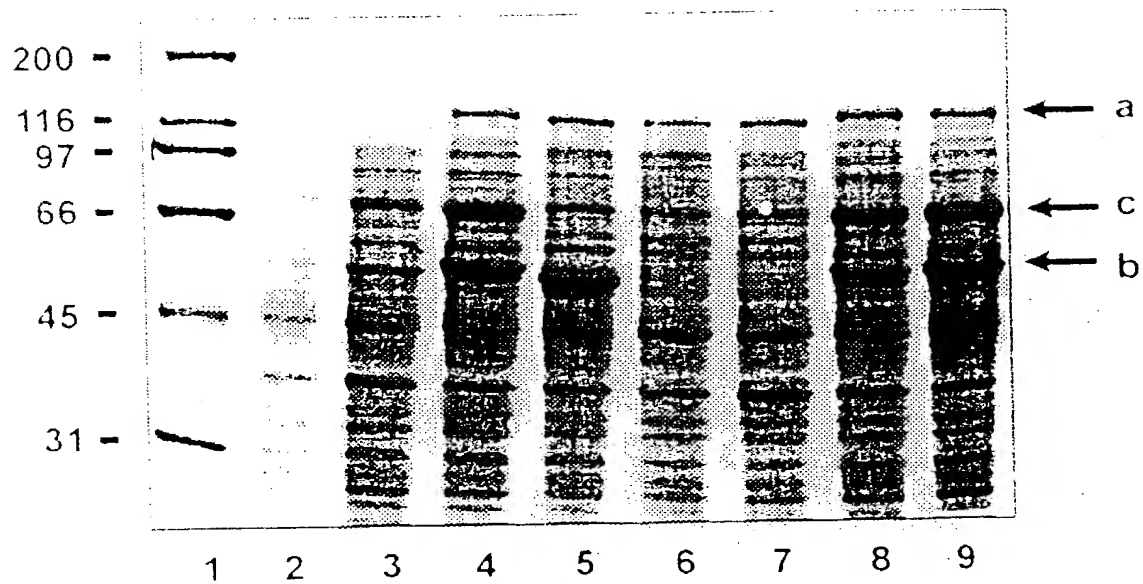
Construction of BK-76-1-1, a *T7 hmw1ABC/cer/KanR* plasmid.

FIG.10

17/235

FIG. 11.



18/235

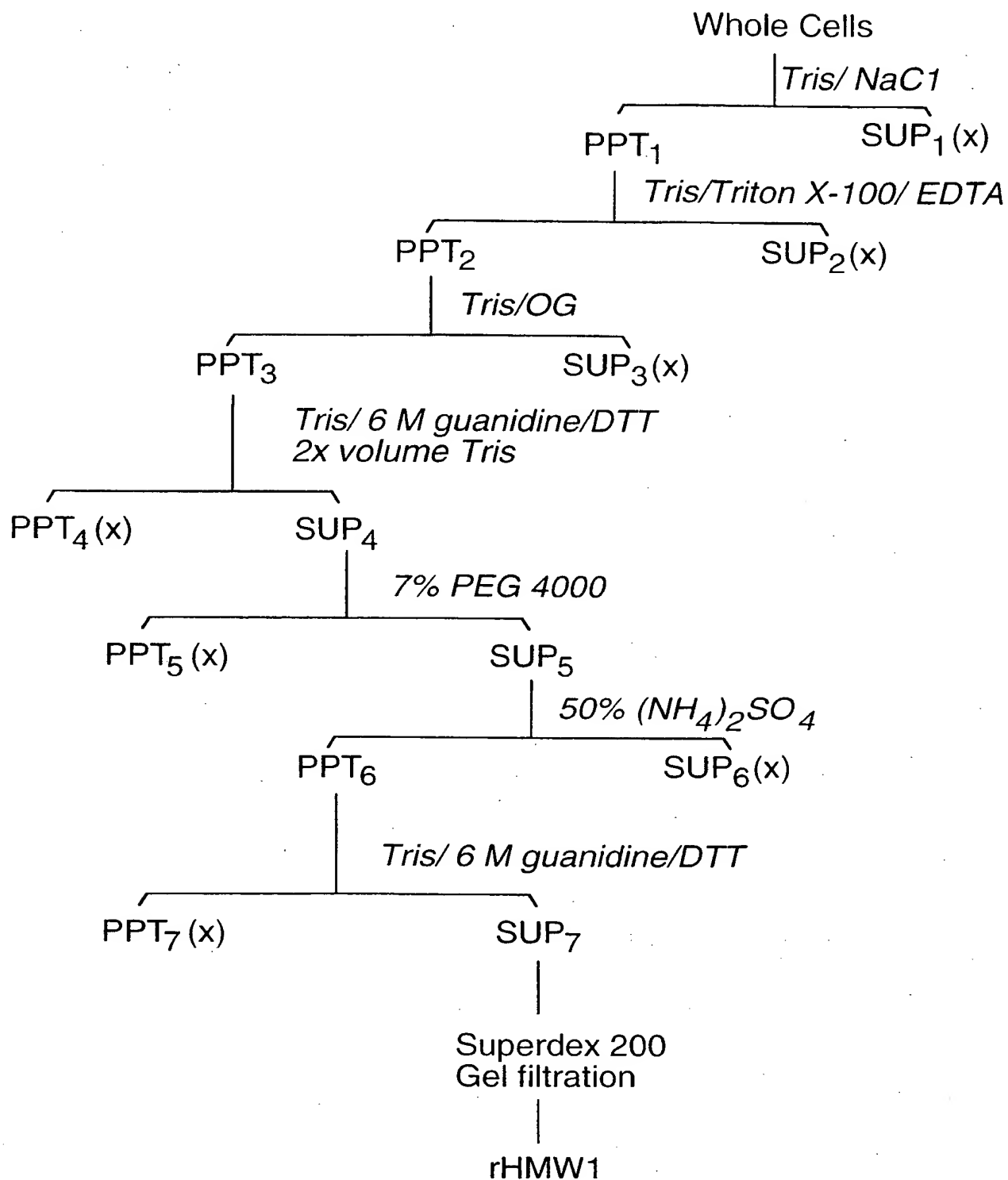
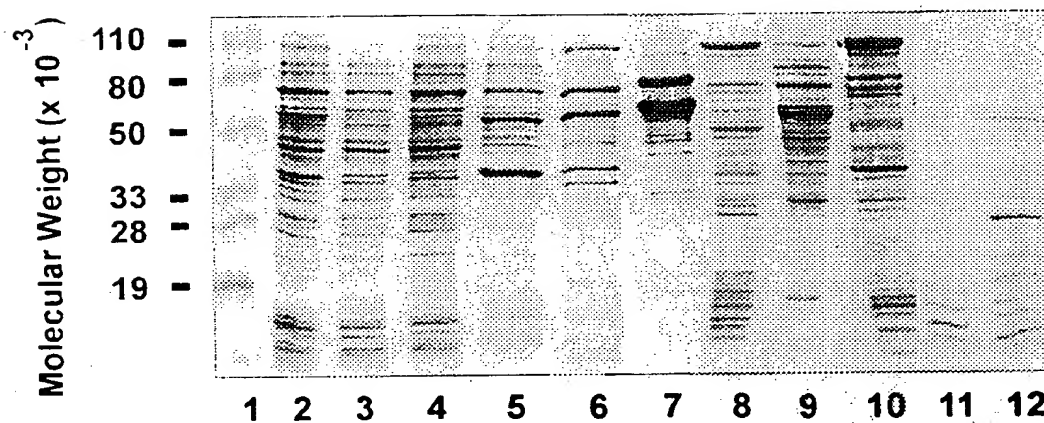
Purification of rHMW1 from *E.coli*

FIG. 12

All centrifugation is at 20,000g, 30min.
(x): to discard

19/235

FIG.13.

Extraction of rHMW1 from *E. coli*

20/235

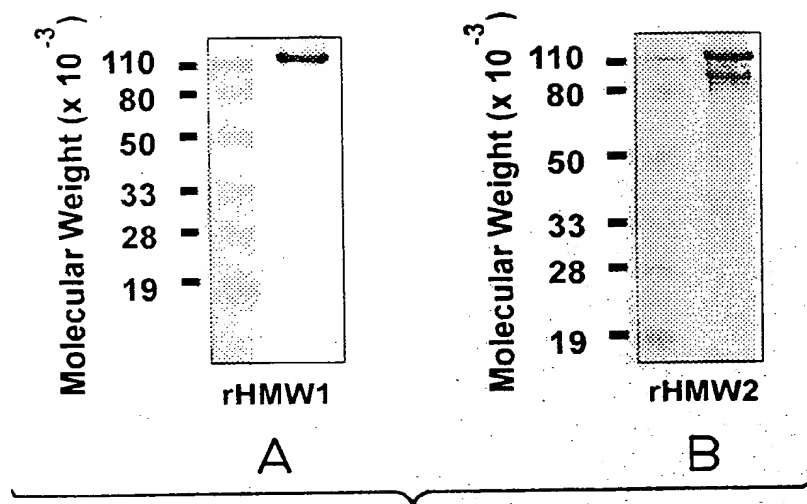
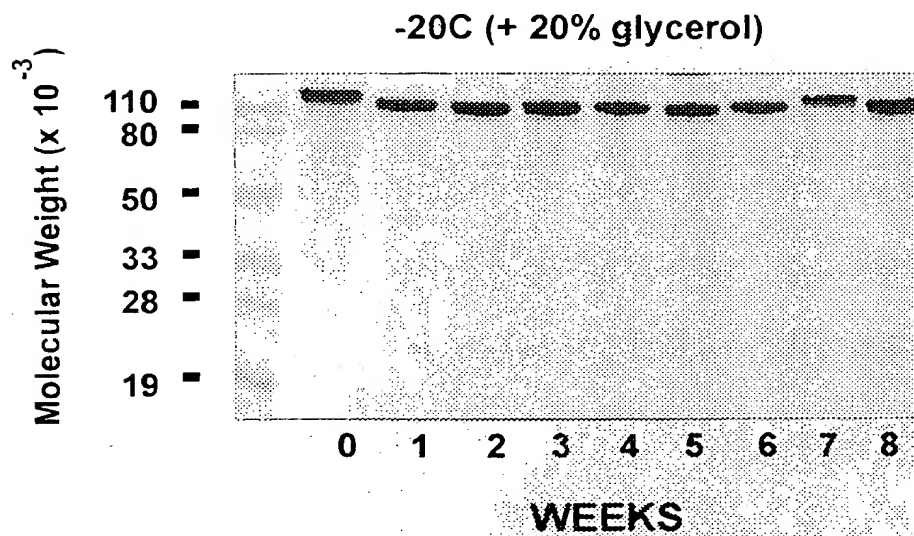
Purification of rHMWs from *E. coli*

FIG.14.

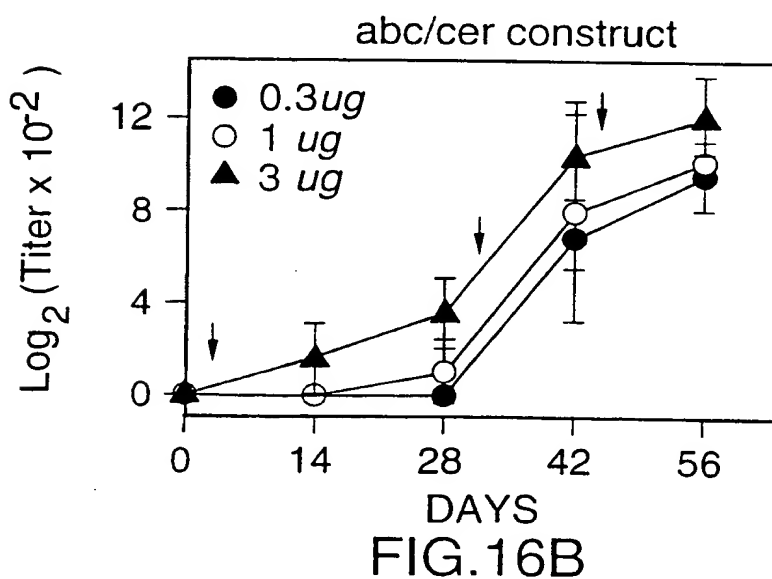
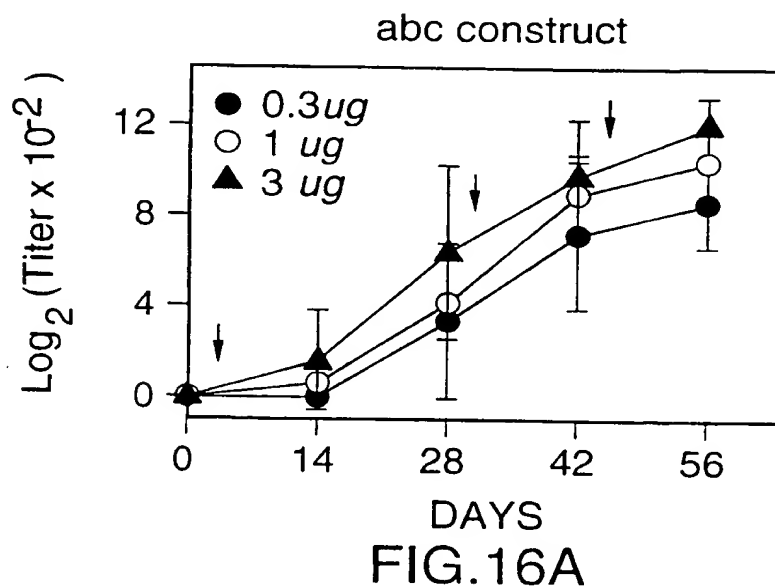
21/235

FIG.15.

The Stability of rHMW1 (abc/cer)

22/ 235

Kinetics of Antibody Response to rHMW1 in Mice



23/235

FIG.17

Oligonucleotides used to PCR amplify *hmmA* genes from NIH strains.

		<u>SEO ID NO</u>	
sense			
	EcoR I ↓		
5'	K E W L L D P	20	
	GCGAATTC	21	
	AAAGAGTGGTTGTAGACCCGGA		5522.SL
	3'		
antisense			
	M K N I K S R L K L	22	
	ATGAAAAATATAAAAGCAGATTAAAC	23	
	TACTTTTATATTTTCGCTAATTTGAGGCTTAAG	24	
	5'		5523.SL
	↑		
	EcoR I		

FIG.18A Joyc hmw1A sequence

LYS GLU TRP LEU LEU ASP PRO ASP ASN VAL ...
 A A G A A T G G T T G T T A G A C C C G G A C A A T G T A T ...
 10 20 30 ...
 ...SER ILE ASN ALA GLY THR SER GLU ARG ASN
 ... C C A T T A A C G C A G G C A C A T C A G A A C G T A A C
 ... 40 50 60
 24/235
 ASP ALA SER PRO THR GLU ASP PHE PRO THR ...
 G A C G C T T C A C C A A C A G A A G A T T T C C C T A C C G ...
 70 80 90 ...
 ...GLY ALA GLY GLY LYS ASP ASN PRO LYS LYS
 ... G A G C A G G A G G A A G G A T A A C C C C A A A A A
 ... 100 110 120
 ASN ALA HIS ASN LYS PRO THR LEU ILE ASN ...
 A A C G C T C A C A C A A A C C G A C A T T A A T A A C A A ...
 130 140 150 ...
 ...THR THR LEU GLU ARG ILE LEU SER GLY ASN
 ...C A A C T C T T G A G C G T A T A T T A G T G G C A A C
 ... 160 170 180
 THR PHE VAL ASN ILE THR ALA ARG LYS ARG ...
 A C C T T G T T A A T A T C A C T G C C A G A A A A G A A ...
 190 200 210 ...

25/235

FIG.18B

...ILE THR VAL ASN SER ASP ILE ASN ILE LYS
 ... TCACAGTTAATAAGTGATATCAACATCAAA
 ... 220 230 240

ASP SER SER HIS LEU ILE LEU TRP SER GLU ...
 GACAGCTCCCATCTAATACTCTGGAGCGAA...
 250 260 270 ...

...ASN ASP ASN SER SER GLY VAL ASP ILE LYS
 ... ATGATAACAGCAGCGCGTTGATATTAA
 ... 280 290 300

GLY ASN ILE THR SER THR GLY GLY SER ...
 GGCAATATCACTTCTACTACTGGCGGAAGCT...
 310 320 330 ...
 ...LEU THR ILE TYR SER SER GLY TRP ILE ASP
 ... TAACTATTACTTCCAGCGCGCTGGATTGAT
 ... 340 350 360

ILE HIS LYS ASN ILE THR LEU ASN SER GLY ...
 ATTCAATAAACAATTACGCTTAATTCAAGGC...
 370 380 390 ...
 ...LEU LEU ASN ILE THR THR LYS GLN GLY ASP
 ... TCTTAACAATTACAACTAAACAAGGAGAT
 ... 400 410 420

FIG.18C

ILE ALA PHE GLU LYS GLY ASN ASN PRO THR ...
 A T C G C C T T C G A A A A G G G A A T A C C C A A C C A ...
 430 440 450 ...
 ... ILE THR GLY GLN GLY THR ILE THR ALA GLY
 ... T T A C A G G T C A A G G G A C T A T T A C C G C A G G C
 460 470 480
 ...
 ASN GLY LYS GLY PHE ARG PHE GLU ASN ALA ...
 A A T G G T A A A G G T T T T A G G T T T G A A A A C G C C T ...
 490 500 510 ...
 ... SER LEU ASN GLY ILE GLY THR GLY LEU LEU
 ... C C C T A A A C G G T A T T G G A A C A G G G T T A C T T
 520 530 540
 ...
 PHE ASN ILE LYS ARG ASP LEU GLY ASN ASN ...
 T T T A A C A T C A A A A G G G A T T T A G G A A A T A A T T ...
 550 560 570 ...
 ... PHE GLN ILE ILE ASN PHE PHE ASN GLY THR
 ... T C C A A A T C A T A A C T T T T T A A C G G A A C T
 580 590 600
 ...
 LEU ASN ILE SER GLY LYS VAL ASN ILE SER ...
 T T A A A T A T T C A G G G A A A G T A A A C A T C T C A A ...
 610 620 630 ...

26/235

FIG.18D

...MET VAL ILE PRO LYS LYS TRP ASP TYR SER
 ... TGGTCAATACCTAAATAATGGATTATAGT
 ... 640 650 660

LYS PHE ARG GLY ARG THR TRP ASN VAL ...
 AATCAGGGGCGAACCCTATTGGAAACGTAA...
 670 680 690 ...

...THR HIS LEU ASN VAL SER GLU SER LYS
 ... CCCATTTAATAATGTTTCCGAAGGCAGTAG
 ... 700 710 720

27/235

PHE ASN LEU THR ILE ASP SER ARG GLY ASP ...
 TTAACTCACTATCGACTCCAGAGGAGATG...
 730 740 750 ...

...ASP THR ALA GLY THR LEU ASN THR PRO TYR
 ... AACTGCAGGCACCCCTTAACACCCCTTAT
 ... 760 770 780

ASN LEU ASN GLY ILE SER PHE ASN LYS ASP ...
 AATTAAACGGTATATCATTCACAAAGACA...
 790 800 810 ...

...THR ILE PHE ASP VAL LYS GLN ASN GLY ALA
 ... CTATCTTTGATGTTAACAACAACGGGGCA
 ... 820 830 840

28/235

FIG.18E

```

VAL  THR  PHE  ASP  ILE  LYS  ALA  PRO  ILE  GLY  ...
GTCACCCTTTGACATCAAGGCGCCAAATAGGGG...
      850
      ...VAL  ASN  ASN  ASN  ARG  ASN  LEU  ASN  TYR  ALA
      ...TAAATAATCGTAATTGAAATTACGCA
      ...
      880
      890
      900

```

```

SER  PHE  ASN  GLY  ASN  ILE  SER  VAL  SER  GLY  ...
TCATTCAATGGAAATAATTTCAGTTTCAGGAG...
      910
      ...GLY  GLY  ASN  VAL  THR  PHE  LYS  LEU  LEU  ALA
      ...GAGGGAATGTCACCTTCAAACTCTCGCC
      ...
      940
      950
      960

```

```

SER  SER  SER  THR  ALA  GLN  THR  PRO  GLY  VAL  ...
TCATCCTCTACCGCTCAAACTCCGGGTGTA...
      970
      ...PHE  ILE  ASN  SER  LYS  HIS  PHE  ASN  ALA  SER
      ...TTATAATTCTAAACACTTTAATGCTTCA
      ...
      1000
      1010
      1020

```

```

GLY  GLY  SER  SER  LEU  GLU  PHE  ARG  THR  GLU  ...
GGAGGGTCGAGTTTAGAATTTAGAACTGAAG...
      1030
      1040
      1050
      ...

```

FIG.18F

WO 00/20609

09/806709

PCT/CA99/00938

29/235

...GLY SER THR LYS VAL GLY PHE LEU ILE ASN
... GCTCAACAAGTCGGCTTCTTGATAAAT
... 1060 1070 1080

ASN ASP LEU THR LEU ASN ALA THR GLY GLY ...
AATGATTTAACCTTAATGCCACCGGAGGTA...
1090 1100 1110 ...

...ASN ILE SER LEU LEU GLN VAL GLU GLY ILE
... ACATATCGCTCTTGCAAGTTGAAGGCATT
... 1120 1130 1140

ASP GLY MET ILE GLY LYS GLY VAL VAL ALA ...
GACGGGATGATTGGTTAAAGGCGTTGTAGCTA...
1150 1160 1170 ...

...LYS LYS ASN ILE THR PHE ALA GLY GLY ASN
... AAAAACAATAACCTTTGCTGGAGGCATA
... 1180 1190 1200

ILE THR PHE GLY SER LYS LYS ALA ILE THR ...
ATCACCTTTGGCTCCAGAAGCAATAACAG...
1210 1220 1230 ...

...GLU ILE GLU GLY ASN ALA THR ILE ASN ASN
... AATCGAAGGCATACTACTAATAAC
... 1240 1250 1260

30/235

FIG.18G

```

ASN  ALA  ASN  VAL  THR  LEU  ILE  GLY  SER  ASP  ...
AACGCTAACGTCACCTCTTATCGGTTTCGGATT...
1270                                     1280
...PHE  ASP  ASN  HIS  GLN  LYS  PRO  LEU  THR  ILE
...TTGACAAACCATCAAAACCTTTTAACTATT
...                                     1300
1310                                     1320

LYS  LYS  ASP  VAL  ILE  ILE  ASN  SER  GLY  ASN  ...
AAAAGATGTCATCATTAATAGCGGCAAC...
1330                                     1340
...LEU  THR  ALA  GLY  GLY  ASN  VAL  ILE  ASN  ILE
...TTACCGCTGGCGGCAATGTTATCAATAATA
...                                     1360
1370                                     1380

ASN  GLY  ASN  LEU  THR  VAL  ASN  ASN  GLY  ALA  ...
ACGGAAATCTTTACCGTTTAAACAATGGCGCCA...
1390                                     1400
...ASN  LEU  LYS  ALA  ILE  THR  ASN  PHE  THR  PHE
...ATCTTAAAGCTATCACAAATTTCACTTTT
...                                     1420
1430                                     1440

ASN  VAL  GLY  GLY  LEU  PHE  ASP  ASN  LYS  GLY  ...
ATGTAGGCGGCTTGTTTTGACCAACAAGGCA...
1450                                     1460
1470 ...

```

31/235

FIG.18H

...ASN SER ASN ILE SER ILE ALA ARG GLY GLY
 ... A T T C A A A T A T C T C C A T T G C C T A G A G G A G G G
 ... 1480 1490 1500

ALA LYS PHE LYS ASP ILE ASN THR SER ...
 G C T A A A T T A A G A T A T C A A T A C C A C C A G T A ...
 1510 1520 1530 ...
 ...SER LEU ASN ILE THR THR ASN SER ASP THR
 ... G C T T A A A T A T T A C C A C C A C T C C G A C A C C
 ... 1540 1550 1560

THR TYR ARG THR ILE ILE GLU GLY ASN ILE ...
 A C T T A C C G T A C C A T T A T A G A A G G T A A T A T A A ...
 1570 1580 1590 ...
 ...THR ASN LYS ALA GLY ASP LEU ASN ILE ILE
 ... C C A A C A A A G C A G G T G A T T T G A A T A T C A T T
 ... 1600 1610 1620

ASP ASN LYS GLY ASN ALA GLU ILE GLN ILE ...
 G A T A A A A G G T A A C G C T G A A A T C C A A A T T G ...
 1630 1640 1650 ...
 ...GLY GLY ASN ILE SER GLN LYS GLU GLY ASN
 ... G C G G C A A T A T C T C G C A A A A G A A G G T A A T
 ... 1660 1670 1680

32/235

FIG.18I

```

LEU THR ILE SER SER ASP LYS ILE ASN ILE ...
CTC ACGATTCTTCCGATAAATAATTAATCA...
1690
1700
...THR ASN GLN ILE THR ILE LYS GLY VAL
...CTAACCCAGATAACAATCAAGAGGGTGT
1720
...
1730
1740

```

```

ASN LYS GLU ASP SER SER THR ALA ...
AATAAGAGGATTCTGATTCAAGCAGGCA...
1750
1760
...ASN ASN ALA ASN LEU THR ILE LYS THR LYS
...ACAATGCTAATCTAACCAATTAAACCAAA
1780
...
1790
1800

```

```

GLU LEU GLN LEU THR GLY ASP LEU ASN ILE ...
GAATTGCAATTAACGGGAGACCTAATAATT...
1810
1820
...SER GLY PHE ASP LYS ALA GLU ILE THR ALA
...CAGGCTTCGATAAAGCAGAAATCACAGCC
1840
...
1850
1860

```

```

LYS GLU GLY ALA ASP LEU ILE ILE GLY ASN ...
AAGAGGGTGCCGATTTAATCATCGGTATA...
1870
1880
1890
...

```

33/235

FIG.18J

...SER ASP ASN ASN ASN ALA ASN ALA LYS
 ...GTGATATAACAACAATGCTAATGCTAAA
 ... 1900 1910 1920

LYS VAL THR PHE ASN GLN VAL LYS ASP SER ...
 AAGTAACCTTTAACCAAGTTAAAGATTCTGA...
 ... 1930 1940 1950 ...

...LYS ILE SER ALA GLY SER HIS ASN VAL THR
 ...AATCTCTGCTGGCAGTCAACAATGTACAA
 ... 1960 1970 1980

LEU ASN SER LYS VAL GLU THR SER ASN GLY ...
 CTAAACAGTAAGTAGAAACCTCTAATGGCA...
 ... 1990 2000 2010 ...

...ASN ASN ASP ALA GLU SER ASN ASN GLY ASP
 ...ATAATGACGCTGAAAGCAATAATGGCGAT
 ... 2020 2030 2040

SER THR SER LEU THR ILE ASN ALA LYS ASN ...
 AGCAACAGCTTAACCTATTATGCAAAATG...
 ... 2050 2060 2070 ...

...VAL THR VAL ASN ASN ASN ILE THR SER HIS
 ...TAACAGTAACAACAATAATTACTCTCAC
 ... 2080 2090 2100

34/235

FIG.18K

```

LYS THR VAL ASN ILE THR ALA SER GLU ASN ...
A A A C A G T A A A T A T C A C T G C G T C A G A A A T G ...
2110                               2120                               2130 ...
...VAL THR THR LYS ALA GLY THR THR ILE ASN
... T T A C C A C C A A A G C G G C A C A C C A T T A A T
...                               2140                               2150                               2160

ALA THR ILE GLY SER VAL GLU VAL THR ALA ...
G C A A C C A T A G G T A G C G T A G A A G T A A C A G C C A ...
2170                               2180                               2190 ...
...LYS THR GLY ASP ILE LYS GLY GLY ILE GLU
... A A A C A G G T G A T A T T A A A G G T G G A A T T G A A
...                               2200                               2210                               2220

SER ASN SER GLY ASN VAL ASN ILE THR ALA ...
T C C A A T T C C G G T A A T G T A A A T A T T A C A G C G A ...
2230                               2240                               2250 ...
...SER GLY ASP THR LEU ASN VAL SER ASN ILE
... G C G G C G A C A C G C T T A A T G T A A G T A A C A T C
...                               2260                               2270                               2280

THR GLY GLN ASN VAL THR VAL ALA ALA ALA ...
A C A G G T C A A A A T G T G A C A G T G C A G C A G C C T ...
2290                               2300                               2310 ...

```

FIG.18L

...SER GLY ALA VAL THR THR THR LYS GLY SER
 ... CAGGTGCCGTAAACACCAACAAGGATCA 2340
 ... 2320 2330

 THR ILE ASN ALA THR THR GLY ASN ALA ASN ...
 ACTATTAAATGCACACACTGGTAAATGCAATAA...
 2350 2360 2370 ...
 ...ILE THR THR LYS THR GLY GLU ILE ASN GLY
 ... TTACACCAACAACAGGTGAAATTAATGGC 2400
 ... 2380 2390

 GLU VAL LYS SER ALA SER GLY ASN VAL ASN ...
 GAGTTAAATCAGCTTCCTGGTAAATGTAATAA...
 2410 2420 2430 ...
 ...ILE THR ALA SER GLY ASN THR LEU ASN VAL
 ... TTACAGCGAGCGGCAATACACTTAATGTA 2460
 ... 2440 2450

 SER ASN ILE THR GLY GLN ASN VAL THR VAL ...
 AGTAACATCACTGGTCAAAATGTAAACAGTA...
 2470 2480 2490 ...
 ...THR ALA ASN SER GLY ALA ILE THR THR THR
 ... CAGCAAACTCAGGTGCCATAACACCAACA 2520
 ... 2500 2510 2520

35/235

36/235

FIG. 18M

GLU GLY SER THR ILE ASN ALA THR THR GLY ...
 GAAGGCTCAACTATTATTAACGGCACAACAGGTG...
 2530 2540 2550 ...
 ...ASP ALA ASN ILE THR THR GLN THR GLY ASN
 ...ATGCAATAATTACAACCCACAACAGGTAAAT
 ... 2560 2570 2580

ILE ASN GLY LYS VAL GLU SER SER SER GLY ...
 ATTAATGGTAAGAATTGAATCCAGTCTCTGTT...
 2590 2600 2610 ...
 ...SER VAL THR LEU ILE ALA THR GLY GLN THR
 ...CTGTGACCGCTTATTGCAACTGGACAACACT
 ... 2620 2630 2640

LEU ALA VAL GLY ASN ILE SER GLY ASP THR ...
 CTTGCTGTAGGTAATAATTTCAGGTGACACTG...
 2650 2660 2670 ...
 ...VAL THR ILE THR ALA ASP LYS GLY LYS LEU
 ...TTACCAATTACTGCGGATAAAGGTAAATAATA
 ... 2680 2690 2700

THR THR GLN THR SER SER LYS ILE ASN GLY ...
 ACCACAACAAGCTCTAAGATTAAACGGA...
 2710 2720 2730 ...

37/235

FIG.18N

...THR LYS SER VAL THR THR SER SER GLN SER
 ... C T A A G A G T G T A A C C A C C T C A A G C C A A T C A
 ... 2740 2750 2760

GLY ASP ILE SER GLY THR ILE SER GLY ASN ...
 G G T G A T A T T A G T G G C A C A A T T C T G G T A A T A ...
 2770 2780 2790 ...

...THR VAL SER VAL SER ALA THR GLY SER LEU
 ... C G G T A A G C G T T A G T G C G A C C G G T A G C T T G
 ... 2800 2810 2820

THR THR GLN ALA GLY SER LYS ILE GLU ALA ...
 A C C A C T C A A G C A G G C T C A A A A A T T G A A G C A A ...
 2830 2840 2850 ...

...LYS THR GLY GLU ALA ASN VAL THR SER ALA
 ... A A A C A G G T G A G G C T A A T G T A A C A A G C G C A
 ... 2860 2870 2880

THR GLY THR ILE GLY GLY THR ILE SER GLY ...
 A C A G G T A C A A T T G G C G G T A C A A T C T C T G G C A ...
 2890 2900 2910 ...

...ASN THR VAL ASN VAL THR ALA ASN THR ASP
 ...A T A C A G T A A A T G T T A C A G C A A A T A C T G A T
 ... 2920 2930 2940

38/235

09/806709

FIG.180

```

ASN  LEU  THR  ILE  LYS  ASP  GLY  ALA  ARG  ILE  ...
A A T T A A C T A T T A A G A T G G C G C A A G A A T T A ...
2950                                     2960
...LYS  ALA  THR  GLY  GLY  ALA  VAL  THR  LEU  THR
... A A G C A A C G G G C G G A G C T G T G A C T T T A A C C
                                     2980
...                                     2990
                                     3000

```

```

ALA  THR  GLY  GLY  THR  LEU  THR  GLU  THR  ...
G C A C A G G A G G T A C T T T A A C C A C C G A A A C A A ...
3010                                     3020
...SER  SER  ASP  ILE  THR  SER  SER  ASN  GLY  GLN
... G T T C T G A T A T T A C C T C A A G C A A T G G T C A G
                                     3040
...                                     3050
                                     3060

```

```

THR  THR  LEU  THR  ALA  LYS  ASP  SER  SER  ILE  ...
A C A C T C T C A C G G C C A A G G A T A G C A G T A T C G ...
3070                                     3080
...ALA  GLY  SER  ILE  ASN  ALA  ALA  ASN  VAL  THR
... C A G G A A G C A T C A A T G C C C G C C A A T G T G A C A
                                     3100
...                                     3110
                                     3120

```

```

LEU  ASN  THR  THR  GLY  THR  LEU  THR  THR  VAL  ...
T T A A T A C C A C A G G C A C T T T A A C T A C T G T G G ...
3130                                     3140
                                     3150

```

FIG.18P

...ALA GLY SER LYS ILE GLU ALA ALA SER GLY
 ... CAGGTTCAAATAATCGAGGCAAGCAGTGGC 3180
 ... 3160 3170

THR LEU VAL ILE ASN ALA LYS ASP ALA GLN ...
 ACCCTGGTTATTATAATGCATAAGATGCTCAGT... 3200
 ... 3190 3210 ...

...LEU ASP GLY ALA ALA LEU GLY ASP ARG THR
 ... TGGACGGCGCGGCATTAGGTGACCGTACA 3240
 ... 3220 3230

GLU VAL ASN VAL THR ASN ALA ASN GLY SER ...
 GAGTAAATGTAACTAACGCATAATGGCTCCG... 3260
 ... 3250 3270 ...
 ...GLY SER VAL ILE ALA THR THR SER SER ARG
 ... GCAGCGTAATCGCGACAACCTCAAGCAGA 3300
 ... 3280 3290

VAL ASN ILE THR GLY ASP LEU ILE THR ILE ...
 GTGAACATCACTGGGGAATTATAACAATAA... 3320
 ... 3310 3330 ...
 ...ASN GLY LEU ASN ILE ILE SER LYS ASN GLY
 ... ATGGATTATAATCATTTCAAAACGGT 3360
 ... 3340 3350

39/235

40/235

FIG.18Q

LYS ASN THR VAL LEU LEU LYS GLY VAL GLU ...
 A A A C A C C G T G C T G T T A A A A G G T G T T G A A A ...
 3370 3380
 ... ILE ASP VAL LYS TYR ILE GLN PRO GLY ILE
 ... T T G A T G T G A A A T A C A T T C A A C C G G C A T A
 3400 3410 3420
 ...

ALA SER VAL TYR GLU VAL ILE GLU ALA LYS ...
 G C G A G C G T A T A T G A A G T A A T T G A A G C A A A C ...
 3430 3440 3450
 ... ARG ALA LEU GLU LYS VAL LYS ASP LEU SER
 ... G C G C T C T T G A G A A A G T G A A A G A T T T A T C T
 3460 3470 3480
 ...

ASP GLU GLU ARG GLU ALA LYS LEU ...
 G A T G A G A A A G A G A A G C A T T A G C T A A G C T T G ...
 3490 3500 3510
 ... GLY VAL SER ALA VAL ARG PHE ILE GLU PRO
 ... G T G T G A G C G C T G T A C G T T T A T T G A G C C A
 3520 3530 3540
 ...

ASN ASN THR ILE THR VAL ASP THR GLN ASN ...
 A A T A A T A C A A T T A C A G T C G A T A C A C A A A T G ...
 3550 3560 3570
 ...

FIG.18R

...GLU PHE ALA THR ARG PRO LEU SER ARG ILE
 ... A A T T G C A A C C A G A C C A T T A A G T C G A A T A
 ... 3580 3590 3600

VAL ILE SER GLU GLY ARG ALA CYS PHE SER ...
 G T G A T T C T G A A G G C A G G C G T G T T C T C A A...
 3610 3620 3630 ...
 ...ASN SER ASP GLY ALA THR VAL CYS VAL ASN
 ... A C A G T G A T G G C G C G A C G G T G T G C G T T A A T
 ... 3640 3650 3660

ILE ALA ASP ASN GLY ARG ***
 A T C G C T G A T A A C G G G C G G T A G
 3670 3680

FIG.19A

Joyc hmw2A sequence



```

    LYS  GLU  TRP  LEU  LEU  ASP  PRO  ASP  ASN  VAL  ...
    A A G A G T G G T T G T T A G A C C C G G A T A A T G T A ...
    10                                     20      30...
        ... SER  ILE  GLU  ASN  PRO  SER  THR  GLU  ARG  ASN
        ... T C C A T T G A A A A T C C C T T C A A C T G A A C G C A A T
        ...                                     40      50      60

    ASP  SER  ASN  GLU  ASP  LEU  GLU  TYR  THR  GLY  ...
    G A T T C C A A T G A A G A C C C T A G A G T A T A C A G G A ...
    70                                     80      90...
        ... THR  GLY  GLU  ASN  ILE  ASN  ASN  PRO  LYS  VAL
        ... A C A G G G G A A A A T A T A A C A C C C T A A G G T A
        ...                                     100      110      120

    ASN  ASN  GLN  SER  LYS  LYS  THR  LEU  THR  SER  ...
    A A T A A T C A G T C T A A A A A A C A C T A A C A A G C ...
    130                                     140      150...
        ... SER  ILE  LEU  GLU  ASN  ILE  LEU  LYS  LYS  GLY
        ... T C A A T C C C T T G A G A A C A T C C C T G A A A A A G G C
        ...                                     160      170      180
  
```

43/235

FIG.19B

SER PHE VAL ASN ILE THR ALA THR ASP ASN ...
 T C T T T G T T A A C A T T A C T G C C A C T G A T A A C ...
 190 200 210...
 ... ILE TYR VAL ASN SER SER ILE ASN ILE GLY
 ... A T C T A C G T T A A T A G C T C T A T C A A C A T C G G A
 220 230 240

ASP SER GLY HIS LEU ILE LEU SER GLY GLY ...
 G A C A G T G G T C A C T T A A T T C T C T C A G G T G G A ...
 250 260 270...
 ... GLY ARG ASN GLY GLY GLY VAL LYS ILE ASN
 ... G G C A G G A A C G G C G G C G G T G T T A A G A T T A A T
 280 290 300

LYS ASN ILE THR SER THR GLY GLY SER LEU ...
 A A A A T A T T A C T T C C A C G G G C G G A A G T T A ...
 310 320 330...
 ... THR ILE ASN SER LYS GLY TRP VAL ASP ILE
 ... A C C A T T A A T T C C A A A G G A T G G G T T G A T A T T
 340 350 360

HIS SER ASN ILE SER LEU GLY THR GLY PHE ...
 C A C T C C A A T A T T C A C T T G G T A C G G G T T T ...
 370 380 390...

44/235

FIG.19C

... LEU ASN ILE THR SER ASN GLY SER VAL ALA
 ... TTGAACATTACCTCTATAATGGTTCCGTTGGCT
 ... 400 410 420

PHE GLU LYS ALA ASP LYS ASP LYS ALA ARG ...
 TTGAGAGGCAAGACAAGATAGGCACGT...
 430 440 450...

... SER ALA ALA ASP ALA GLN ILE VAL ALA GLN
 ... AGCGCGGCAGATGCTCAATAATTGTCGCCACA
 ... 460 470 480

GLY ILE ILE ASN LEU THR GLY GLU ASN LYS ...
 GGCAATCAATAACCTCACAGGGGAACAACA...
 490 500 510...

... THR PHE ARG LEU ASN ASN VAL SER LEU ASN
 ... ACCTTTAGGCTTAACAATAATGTCCTTTAAT
 ... 520 530 540

GLY VAL GLY GLN GLY LEU SER ILE THR SER ...
 GGAGTGGGTCAAGGTCCTATCCATCACGTC...
 550 560 570...

... ASN VAL GLY ASN GLN THR HIS LYS PHE ASP
 ... ATGTGGGCAATCAACAATCAATAATTCGAT
 ... 580 590 600

45/235

FIG.19D

```

GLY  GLU  ILE  ASN  ILE  THR  GLY  ASN  VAL  THR ...
GGTGAATTAACATAACTGGAAATGTACA...
610                                     620
... ILE  ASN  GLN  THR  ALA  PRO  ALA  THR  THR  ALA
... ATTAATCAAACTGCCACCTGCCGACACCGCA
640                                     650
...                                     660

TYR  TRP  ASN  PHE  SER  TYR  ASP  SER  TYR  TRP ...
TATGGAAATTTAGCTACGATTCTATATGG...
670                                     680
... ASN  VAL  SER  THR  LEU  ASN  VAL  GLN  LYS  ASN
... ACGTCAGTACTCTTAACGTACAAACAAAC
700                                     710
...                                     720

SER  SER  PHE  THR  PHE  ILE  LYS  ARG  THR  GLU ...
TCAGCTTTACCTTTTATTATAGCGCACTGA...
730                                     740
... SER  ASN  ARG  PHE  GLY  PRO  THR  THR  PRO  LEU
... AGTAATCGCTTTTGGCCCAACACCACTT
760                                     770
...                                     780

ARG  SER  SER  GLY  GLY  VAL  PHE  PHE  ASN  GLY ...
CGAGCTCCGGAGGGGTATTCTTTACGGC...
790                                     800
...                                     810...

```

46/235

FIG.19E

... THR ASN GLY ASN MET VAL LEU ASN VAL GLY
 ... ACGAATGGCAACATGGTGCTTAACGTCGGA 840
 ... 820

THR ASN SER ARG VAL LEU PHE ASN LEU LYS ...
 ACTAATTCGAGAGTTTGTTTATAATTGAAG...
 ... 860
 ... 870...

... PRO ASN GLU ASN THR ASN ASN SER LYS PRO
 ... CCAATGAGAAATACAAACAACAGCAAGCCT 900
 ... 880

LEU PRO LEU GLN PHE ASN ALA ASN ILE THR ...
 TACCGCTTCAATTAAACGCCAATAATACA...
 ... 910
 ... 920
 ... 930...

... ALA ILE GLY GLY SER VAL SER PHE ASP
 ... GCCATTGGTGGAGGCTCTGTGTCCTTTGAT 960
 ... 940

ILE HIS ALA ASN HIS SER GLY ARG GLY ALA ...
 ATACGCCCAATCATTCCTCCGGCAGAGGGCT...
 ... 970
 ... 980
 ... 990...

... GLU LEU LYS MET ASN THR ILE ASN ILE SER
 ... GAATTAATAATGAACACCAATTAATATCTCT 1020
 ... 1000
 ... 1010

47/235

FIG.19F

```

ASP  GLY  THR  SER  LEU  THR  LEU  GLN  SER  HIS  ...
GACGGCACCAGCCTCACCCCTACAAATCCCAT...
1030
...  VAL  ARG  LYS  ASP  SER  ALA  PHE  ILE  ILE  SER
...  GTTCGCAAGAATAGTGCTTTTATAATCAGT
1040
...
1050...
1060
1070
1080

LYS  ASP  LEU  THR  ILE  ASN  ALA  THR  GLY  SER  ...
AAGATTTAACAATAAACGCAACCGGTTCA...
1090
...  ASN  PHE  THR  LEU  GLU  GLN  SER  PRO  ASP  SER
...  AATTTACTCTTGAGCAATCACCAAGACAGT
1100
...
1110...
1120
1130
1140

PHE  THR  ASP  LYS  TYR  PRO  GLY  ARG  ALA  ILE  ...
TTTACTGACAAATAACCCCGGAAGAGCTATT...
1150
...  SER  SER  THR  LYS  ASN  ILE  THR  ILE  SER  GLY
...  AGTTCAACTAAATAATAACCATCTCTCAGGT
1160
...
1170...
1180
1190
1200

GLY  ASN  VAL  SER  LEU  GLY  GLY  GLN  ASN  SER  ...
GGCAACGTCCTCTCTGGGTGGGCAAAATTC A...
1210
1220
1230...

```

48/235

FIG.19G

... SER SER ASP ILE LYS GLY ASN ILE THR ILE
 ... AGCAGTGACATCAAGGGAATAATTACCATC
 ... 1240 1250 1260

LYS SER SER THR ASN VAL LEU LYS ALA ...
 AAAGCTCAACAATACTGTAAGCC...
 1270 1280 1290...

... HIS ASN SER PRO ARG ASP PHE ALA SER ARG
 ... CATACAGCCCTCGCGGACTTTGCTTCCAGA
 ... 1300 1310 1320

THR LEU THR LEU GLY ASN LEU VAL GLU ...
 ACCTTAACCCCTTGGCAACTTGAAATGTGA...
 1330 1340 1350...

... GLY ASN LEU THR LEU THR GLY SER VAL ALA
 ... GGAAATTTAACCCCTAACCGGCTCAGTTGGC
 ... 1360 1370 1380

ASP ILE LYS GLY ASN LEU SER ILE LEU ASN ...
 GATATTAAAGGTAACCTTTCCATTCTTAAC...
 1390 1400 1410...

... ASP ALA THR PHE LYS GLY GLU THR SER GLU
 ... GATGCTACTTTTAAAGGAGAGACCAAGTGAA
 ... 1420 1430 1440

49/235

FIG.19H

ASN LEU ASN ILE THR GLY ASN PHE THR ASN ...
 AACCTAAACATCACCGGCAACTTCACCAAT...
 1450 1460 1470...
 ... ASN GLY THR ALA ASP ILE ASN ILE LYS GLN
 ... AATGGCACCGCCGACATTAAATATAAACA
 ... 1480 1490 1500

GLY VAL VAL ASN ILE GLN GLY ASN ILE THR ...
 GGGTGGTAACATCCCAAGGTAAATATACC...
 1510 1520 1530...
 ... ASN LYS GLY GLY LEU ASN ILE THR ASN
 ... AATAAGGTGGTTTAAACATTACCACTAAT
 ... 1540 1550 1560

ALA GLN ASN ASN GLN LYS THR ILE ILE ASN ...
 GCCCAAAACAATCAAAACCAATTATTAAAC...
 1570 1580 1590...
 ... GLY ASN ILE THR ASN GLU GLY GLY ASP LEU
 ... GGAAATATACTAACGAAGGCGGAGATTATA
 ... 1600 1610 1620

ASN ILE LYS ASP SER ASN ASN ALA GLU ...
 AACATCAAGGATAGTAACAATAATGCTGA...
 1630 1640 1650...

09/806709

50/235

FIG.19I

... ILE GIN ILE GLY GLY ASN ILE SER GLN LYS
 ... ATCCAAATTGGCGGCAATATCTCGCAAAA
 ... 1660 1670 1680

LYS GLY ASN LEU THR ILE SER SER ASP LYS ...
 AAGGCAATCTCACAAATTCTCTGTGATAA...
 1690 1700 1710...
 ... ILE ASN ILE THR LYS LYS ILE THR ILE LYS
 ... ATCAATAATTACCACAGAGATACCAATCAAA
 ... 1720 1730 1740

ALA GLY VAL ASP GLU GLY GLY SER ASP SER ...
 GCAGGCGTTGATGAGGTGGTTCTGTCTCA...
 1750 1760 1770...
 ... SER PRO ALA SER ASN ALA ASN LEU THR ILE
 ... AGCCAGCAAGTAATGCTAATCTAACCAATT
 ... 1780 1790 1800

LYS THR LYS THR LEU GLU LEU THR GLY ASP ...
 AAACCAAAACGCTAGAAATTACAGGAGAC...
 1810 1820 1830...
 ... LEU ASN ILE SER GLY PHE ASN LYS ALA GLU
 ... CTAATAATTTCAGGCTTTATAAGCAGAA
 ... 1840 1850 1860

51/235

FIG.19J

```

ILE THR ALA LYS ASN GLY ASN ASP LEU THR ...
ATTACAGCTAAATAAGGCAACGATTTAAC T...
1870
1880
... ILE GLY LYS ALA SER ASP GLY ASN ALA ASN
... ATTGGCAAGGCTAGTGATGGTAATGCTAAT
1900
1910
1920

```

```

ALA LYS LYS VAL THR PHE ASP LYS VAL LYS ...
GCTAAATAAGTGACTTTTGACAAAGGTTAA A...
1930
1940
... ASP SER LYS ILE SER ALA ASN GLY HIS ASN
... GATTCAAATACTCTCAGCTAACGGTCACAAT
1960
1970
1980

```

```

VAL THR LEU ASN SER LYS VAL GLU THR SER ...
GTAACTAAATAAGCAATAAGTGGAACGTC T...
1990
2000
... ASN SER ASP SER SER ALA ASP ASP SER ASN
... AATAGTGATAGTAGTGCTGATGATAGTAAT
2020
2030
2040

```

```

ASP ASN ASN THR GLY LEU THR ILE SER ALA ...
GATAACAACACTGGTTTAACCAATTCCGCA...
2050
2060
2070...

```

FIG.19K

... LYS ASP VAL THR VAL ASN ASN ASP VAL THR
 ... A A G A T G T A A C A G T A A C A A T G A C G T C A C C
 ... 2080 2090 2100

SER HIS LYS THR ILE ASN ILE SER ALA THR ...
 T C C C A C A A G A C A A T A A T A T C T C T G C C A C A ...
 2110 2120 2130...

... THR GLY ASN VAL THR THR LYS GLU SER THR
 ... A C A G G A A A T G T A A C A C C A A A G A A A G C A C A
 ... 2140 2150 2160

THR ILE ASN ALA ALA THR GLY SER VAL GLU ...
 A C C A T T A A T G C G G C C A C A G G T A G C G T G G A A ...
 2170 2180 2190...

... VAL THR ALA LYS THR GLY ASP ILE SER GLY
 ... G T A A C T G C T A A A C A G G C G A T A T T A G T G G C
 ... 2200 2210 2220

THR ILE SER GLY ASN THR VAL ASN VAL THR ...
 A C A A T T C T G G T A A T A C A G T A A A T G T T A C A ...
 2230 2240 2250...

... ALA THR ASP SER LEU THR THR GLN ALA SER
 ... G C A A C T G A T A G C T T A C C C A C C C A G C A A G C
 ... 2260 2270 2280

53/235

FIG.19L

SER SER ILE THR SER SER ASN GLY GLN THR ...
 TCTAGCATTAACCTCAAGTAATGGTCAGACA...
 2290 2300 2310...
 ... THR LEU THR ALA LYS ASN GLY SER ILE ALA
 ... ACTCTTACAGCCCAAGAAATGGCAGTATCGCA
 2320 2330 2340

GLY SER ILE ASP ALA ALA ASN VAL THR LEU ...
 GGAGTATTGATGCCGCTAATGTGACATT A...
 2350 2360 2370...
 ... ASN THR THR GLY THR LEU THR VAL ALA
 ... AATACCACAGGCACCTTAAC TACTGTAGCG
 2380 2390 2400

GLY SER ASN ILE LYS ALA THR SER GLY THR ...
 GGTTCAAACATTAAAGGCAACCAAGTGGA CT...
 2410 2420 2430...
 ... LEU ALA ILE ASN ALA LYS ASP ALA LYS LEU
 ... TTAGCTATTTAACGCCA AAGATGCTAAGTTA
 2440 2450 2460

ASP GLY THR ALA SER GLY ASP ARG THR VAL ...
 GATGGTACTGCATCAGGTGACCGCACAGTA...
 2470 2480 2490...

54/235

FIG.19M

... VAL ASN ALA THR ASN ALA SER GLY SER GLY
 ... GTAAATGCAACTAACGCAGTGGCTCTGGT
 ... 2500 2510 2520

SER VAL THR ALA ALA THR SER SER ASN VAL ...
 AGTGTGAC TGGGCAACCTCAAGTAACGTG...
 2530 2540 2550...
 ... ASN ILE THR GLY ASP LEU SER THR ILE ASN
 ... AATATCACTGGAGATTTAAGCACATAAAT
 ... 2560 2570 2580

GLY LEU ASN ILE ILE SER LYS ASN GLY LYS ...
 GGATTAATAATCATTTCCGAAATAATGGTAA...
 2590 2600 2610...
 ... ASN THR VAL VAL LEU LYS GLY ALA GLU ILE
 ... AACACCGTAGTGTTAAAGGTGCTGAATTT
 ... 2620 2630 2640

ASP VAL LYS TYR ILE GLN PRO GLY VAL ALA ...
 GATGTGAATAATATTCAACCAAGGTGTAGCA...
 2650 2660 2670...
 ... SER ALA ASN GLU VAL ILE GLU ALA LYS ARG
 ... AGTGCGAATGAGGTATTGTGAGCGAGCGT
 ... 2680 2690 2700

55/ 235

FIG.19N

ALA LEU GLU LYS VAL LYS ASP LEU SER ASP ...
 GCCCTTGAAAGTAAGTAAAGATTATCTGAT...
 2710 2720 2730...
 ... GLU GLU ARG GLU THR LEU ALA LYS LEU GLY
 ... GAAGAAAGAGAAACATTAGCTAAACTTGGT
 ... 2740 2750 2760

VAL SER ALA VAL ARG PHE VAL GLU PRO ASN ...
 GTAGTGCTGTACGTTTGTGTGAGCCAAAT...
 2770 2780 2790...
 ... ASN THR ILE THR VAL ASN THR GLN ASN GLU
 ... AATACAATTACAGTCAATACACAATAAGAA
 ... 2800 2810 2820

PHE THR THR ARG PRO SER SER GLN VAL THR ...
 TTACAACAGACCGTCAAGTCAAGTGACA...
 2830 2840 2850...
 ... ILE SER GLU ASP LYS ALA CYS PHE SER SER
 ... ATTCTGAAGACAAAGCGGTGTTCTCTCAGT
 ... 2860 2870 2880

56/235

FIG.190

```
GLY ASN GLY ALA VAL CYS THR ASN VAL ...  
GGTAA TGGTGCAGCAGTATGTACTAATGTT...  
2890 2900 2910...  
... THR ASP ARG GLN ***  
... ACTGACGATAGACAGTAA  
... 2920
```

57/ 235

FIG.20A

K1 *hmm1A* sequence

```

    LYS  GLU  TRP  LEU  LEU  ASP  PRO  ASP  ASN  VAL ...
    A A G A G T G G T T A G A C C C G G A T A A T G T A ...
    10
    ... SER  ILE  ASN  ALA  PRO  ALA  LEU  GLY  ARG  THR
    ... T C T A T T A A T G C A C C C G C A C T T G G A C G T A C T
    ...
    20
    30...
    40
    50
    60

    GLU  SER  THR  PRO  ASN  ASN  ASN  GLU  TYR  ASP ...
    G A G A G T A C C C C A A T A C A A T G A G T A C G A C ...
    70
    ... SER  PRO  ASN  GLN  ILE  ASN  TYR  LYS  ASN  LYS
    ... T C G C C A A A T C A A A T T A C T A T A A A C A A A
    ...
    80
    90...
    100
    110
    120

    PRO  SER  LEU  SER  THR  LEU  THR  ASN  THR  THR ...
    C C A T C C C T A A G T A C A C T A A C A A A C A C A A ...
    130
    140
    150...
    ... LEU  GLU  ARG  ILE  LEU  LYS  ARG  ASN  THR  SER
    ... C T T G A G A G A A T A T T A A A A G A A C A C C T C T
    ...
    160
    170
    180
  
```

58 / 235

FIG.20B

VAL ASN ILE THR ALA THR LYS THR ILE THR ...
 GTTAAATATCAGCTGCCCACCAAAACAATCACA...
 190 200 210...
 ... VAL ASN SER ASP ILE ASN ILE GLY ASP SER
 ... GTTAAATAGTGATATCAATATTGGAGACAGC
 220 230 240
 ...

SER HIS LEU THR LEU TRP SER GLU GLY GLN ...
 TCCCACTTAACCCCTTTGGAGTGAGGGTCAG...
 250 260 270...
 ... GLY ARG GLY GLY VAL ASN VAL THR GLY ASN
 ... GGGAGAGCGCGCGTTAATGTTACAGGCAAT
 280 290 300
 ...

ILE THR SER THR THR ASN GLY ASN LEU THR ...
 ATTACTTCTACTACCAACGGAAACTTAACC...
 310 320 330...
 ... ILE TYR SER GLY TRP VAL ASP VAL HIS
 ... ATTACTCTGGCGGATGGGTTGATGTTCAAT
 340 350 360
 ...

LYS ASN ILE THR LEU LYS SER GLY TYR LEU ...
 AAAACAATTACACTTAATAATCAGGGTACTTA...
 370 380 390...

59/235

FIG.20C

... ASN ILE THR THR LYS GLN GLY ASP ILE ALA
 ... AACATTACAACCTAAACAAGGAGACATCGCC
 ... 400 410 420

PHE GLU ASP LYS PRO GLY LEU SER ASN LEU ...
 TTCGAAGACAAACCAGGCTGAGCAACCTA...
 430 440 450...
 ... THR ILE THR ALA LYS GLY THR ILE ALA VAL
 ... ACCATTACAGCTAAAGGACCAATTGCCGTG
 ... 460 470 480

ASN ASN LYS LYS GLY PHE ARG PHE ASP ASN ...
 AACACAAGAAAGGCTTTAGGTTTGATAAT...
 490 500 510...
 ... VAL THR LEU ASN GLY THR GLY GLY LEU
 ... GTCACTCTAAATGGAACGGAGGAGGCTC
 ... 520 530 540

SER PHE LYS TYR ILE GLU THR GLY ASN ARG ...
 TCTTTTAAATACATCGAAACCGGAATAGA...
 550 560 570...
 ... ASP SER ASN PHE GLU THR HIS PHE ARG GLY
 ... GATAGCAATTTCGAAACCCATTATTAGAGA
 ... 580 590 600

60/235

FIG.20D

ARG LEU ASN ILE SER GLY LYS VAL ASP ILE ...
 A G A T T A A T A T T C A G G G A A A G T A G A T A T C ...
 610 620 630...
 ... LEU MET GLN ALA ARG GLN GLU ASN TRP ASN
 ... T T A A T G C A A G C A A G G C A G G A G A A C T G G A A C
 ... 640 650 660

ARG ARG HIS TRP GLY ARG SER HIS TRP ASN ...
 C G C A G A C A C T G G G A C G C T C C C A C T G G A A T ...
 670 680 690...
 ... VAL THR ARG LEU ASN VAL SER GLU ASN SER
 ... G T A A C C C G A T T G A A C G T T C T G A A A C A G T
 ... 700 710 720

TYR PHE ASN VAL THR ILE ASP SER SER GLY ...
 T A T T T A A C G T C A C T A T T G A T A G C A G T G G C ...
 730 740 750...
 ... SER ALA SER SER PRO GLY ALA GLY PRO LEU
 ... A G T G C C T C T T C C C C T G G C G C T G G C C C T C T G
 ... 760 770 780

ASN ALA GLN SER GLY LEU ASN GLY ILE SER ...
 A A T G C C C A A T C G G G T T T A A A T G G C A T A T C G ...
 790 800 810...

61/ 235

FIG.20E

... PHE ASN ASN ASP THR VAL PHE ASN ILE ALA
 ... TTTAATAATGACACTGTTTATAATAATTGCA
 ... 820 830 840

ALA SER SER ALA VAL ASN PHE ASN ILE LYS ...
 GCAAGTTCGGCGGTTAACTTTAACAATCAAA...
 ... 850 860 870...
 ... PRO PRO ILE VAL ASP LYS VAL THR ASN GLY
 ... CCACCAATAGTAGACAAAGTAACCAACGGG
 ... 880 890 900

ASN HIS THR LEU PHE LYS GLY ASN ILE SER ...
 ...
 AATCACACATTATTCAAAGGGAATATTCA...
 ... 910 920 930...
 ... VAL LEU GLY GLY GLY MET SER THR PHE ILE
 ... GLY GLY ASP VAL ASN PHE HIS P
 ... GTTTTAGGGGGGAGTGTCAACTTCAAT
 ... 940 950 960

HE ASN ALA SER SER SER ASN TYR GLN THR ...
 TTAACGCCCTCCAGCAACTACCAACTT...
 ... 970 980 990...

62/ 235

FIG.20F

...TYR GLY VAL ILE ILE GLU SER GLN ASN PHE S
 ... A T G G C G T G A T T A T A G A G T C A C A A A C T T T A
 ... 1000 1010 1020

ER ALA SER GLY GLY SER LEU LYS PHE ...
 G T G C C T C A G G A G G T C A A G C T T A A A T T C A ...
 1030 1040 1050...
 ...LYS SER GLY GLY SER THR HIS ALA ALA PHE T
 ... A A G C G A A G G T T C G A C A C A C G C C T T T A
 ... 1060 1070 1080

HR ILE LYS ASN ASP LEU ILE LEU ASN ALA ...
 C A A T A A A A A T G A T T T A A A T G C C A ...
 1090 1100 1110...
 ...THR GLY GLY ASN ILE SER LEU ASN GLN VAL A
 ... C T G G G G G C A A T A T A T C A T T G A A C C A A G T T G
 ... 1120 1130 1140

LA GLY ILE ASP SER ASN LEU LYS LYS SER ...
 C A G G T A T T G A T A G T A A T C T C A A A A A G C C ...
 1150 1160 1170...
 ...LEU ILE ALA ASN LYS ASN ILE THR PHE GLU G
 ... T T A T A G C C A A T A A A A C A T A A C C T T T G A A G
 ... 1180 1190 1200

63/235

FIG.20G

```

LY  GLY  ASN  ILE  THR  LEU  ALA  ALA  ASP  LYS  ...
GGGGCAATATCACTTGGAGCCGATAA...
1210                                     1220
...LYS  PRO  ILE  GLU  ILE  LYS  GLY  ASN  ILE  THR  V
...AACCAATAGAAATCAAGGTAAATATTACTG
...                                     1240
                                     1250
                                     1260

AL  LYS  GLU  GLY  ALA  ASN  VAL  THR  LEU  ARG  ...
TTAAGAGAGGAGCCAAATGTCACCTTCGTA...
1270                                     1280
...SER  ALA  ASN  TYR  GLY  ASN  ASP  LYS  SER  ALA  L
...GGCGGAATTATGGTAATGACAAATCAGCTT
...                                     1300
                                     1310
                                     1320

EU  SER  ILE  ARG  GLY  ASN  VAL  THR  ASN  LYS  ...
TAGTATAAGAGGAATAATGTCATAAAG...
1330                                     1340
...GLY  ASN  LEU  THR  VAL  THR  GLY  SER  ALA  ILE  A
...GCAATCTCACCGTTACCGGCTCCGCTATCA
...                                     1360
                                     1370
                                     1380

SN  ILE  GLU  LYS  ASN  LEU  THR  VAL  GLU  GLY  ...
ATATAGAAAAAATACTTACCGTTGAGGTA...
1390                                     1400
                                     1410

```

64/235

FIG.20H

...SER ALA LYS PHE LEU ALA ASN PRO ASN TYR S
 ... GTGCTAAGTTT TAGCTAATCCAAATTAACA
 ... 1420 1430 1440

ER PHE ASN VAL SER GLY LEU PHE ASP ASN ...
 GCTTTAACGTATCCGGCTATTTGACACAC...
 1450 1460 1470...
 ...GLN GLY LYS SER ASN ILE SER ILE ALA LYS G
 ... AAGGCAAGTCACACATTTCCATCGCTAAGG
 ... 1480 1490 1500

LY GLY ALA ILE PHE LYS ASP ILE GLU ASN ...
 GAGGAGCTATTTTAAAGATATCGAGATA...
 1510 1520 1530...
 ...THR GLY SER LEU ASN ILE THR THR LYS SER A
 ... CTGGCAGTCTGAATAATTACCACTAATAATCCG
 ... 1540 1550 1560

SP SER ASN HIS HIS THR ILE ILE LYS GLY ...
 ACTCCAAACCACCTACTATTAATAAGGTA...
 1570 1580 1590...
 ...ASN ILE THR ASN ARG LYS GLY ASP LEU ASN I
 ... ATATACTAACAGAAAGGTGATTATAATA
 ... 1600 1610 1620

65/235

FIG.20I

LE THR ASN ASN GLY ASP ASN THR GLU ILE ...
 TCACGAAATAATGGTGATAATAC TGAAATCC...
 1630 1640 1650...
 ...GLN ILE GLY GLY ASN ILE SER GIN LYS GLU G
 ... AAT TGGCGGCAATA TCTCGCAAAAGAG
 ... 1660 1670 1680

LY ASN LEU THR ILE SER SER ASP LYS VAL ...
 GCATCTCACAAATTTCTTCTTGATAAAGTCA...
 1690 1700 1710...
 ...ASN ILE THR GLU ARG ILE THR ILE LYS ALA G
 ... A TAT TACCGAGCGGATAACAATCAAGCAG
 ... 1720 1730 1740

LY VAL ASN GLY ASP ASN SER ASP SER ASN ...
 GCGTTAATGGGGAATACTCTGTGATTCAATG...
 1750 1760 1770...
 ...GLU ALA THR SER ALA ASN LEU THR ILE LYS T
 ... AGGCAACAAGTGCTAACCTAACCATTAATAA
 ... 1780 1790 1800

HR LYS GLU LEU LYS LEU THR ASN ASP LEU ...
 CCAAGAGTTAAATAATAACAAACGACCTAA...
 1810 1820 1830...

66/235

FIG.20J

```

...ASN ILE SER GLY PHE ASN LYS ALA GLU ILE T
... A T A T T C A G G T T T A A T A A G C A G A A T T A
... 1840 1850 1860

HR ALA LYS ASP ASN SER LEU THR ILE ...
C A G C T A A A G A T A A C A G T A A T T A C T A T T G...
1870 1880 1890...
...GLY ASP ASN SER ASP ALA GLY ASN THR ASP A
... G C G A T A A C A G T G A C G C T G G C A A T A C T G A C G
1900 1910 1920
...

LA LYS LYS VAL THR PHE SER ASN VAL LYS ...
C T A A A A A G T A A C C T T T A G C A A T G T T A A A G...
1930 1940 1950...
...ASP SER LYS ILE SER ALA SER ASP HIS ASN V
... A T T C A A A A T C T C T G C T A G C G A C C A T A A T G
1960 1970 1980
...

AL THR LEU ASN SER LYS VAL GLU THR SER ...
T A A C G C T A A A C A G C A A A G T G G A A A C A T C T G...
1990 2000 2010...
...GLY ASP THR ASP SER THR GLU ASP GLY GLY A
... G C G A T A C T G A C A G C A C T G A A G A T G G C G G C A
2020 2030 2040
...
```

67/235

FIG.20K

SN ASN ASN THR GLY LEU THR ILE THR ALA ...
 A C A A T A A C A C C G G C T T A A C T A T T A C T G C A A ...
 2050 2060 2070...
 ...LYS ASN VAL THR VAL ASN ASN ILE THR S
 ... A A A T G T A A C A G T A A A C A C A A T A T T A C T T
 ... 2080 2090 2100

ER HIS LYS THR VAL ASN ILE THR ALA SER ...
 C T C A C A A A C A G T A A A T A T C A C T G C G T C A G ...
 2110 2120 2130...
 ...GLU ASN VAL THR THR LYS ALA GLY THR THR I
 ... A A A T G T T A C C A C C A A A G C G G C A C A C C A
 ... 2140 2150 2160

LE ASN ALA THR THR GLY SER VAL GLU VAL ...
 T T A A T G C A A C C A C A G G T A G C G T A G A A G T A A ...
 2170 2180 2190...
 ...THR ALA LYS THR GLY ASP ILE LYS GLY GLY I
 ... C A G C C A A A C A G G T G A T A T T A A A G G T G G A A
 ... 2200 2210 2220

LE GLU SER ASN SER GLY ASN VAL ASN ILE ...
 T T G A A T C C A A T T C C G G T A A T G T A A T A T T A ...
 2230 2240 2250...

68/235

FIG.20L

...THR ALA SER GLY ASP THR LEU ASN VAL SER A
 ... CAGCGAGCGCGACACGCTTAATGTAGTA 2280
 ... 2260 2270

SN ILE THR GLY GLN ASN VAL THR VAL ALA ...
 ACATCAGGTCAAAATGTGACAGTGCGAG...
 2290 2300 2310...
 ...ALA ALA SER GLY ALA VAL THR THR LYS G
 ... CAGCCTCAGGTGCCGTAAACAACCAACAAG 2340
 ... 2320 2330

LY SER THR ILE ASN ALA THR THR GLY ASN ...
 GATCAACTATTATGCAACACTGGTAATG...
 2350 2360 2370...
 ...ALA ASN ILE THR THR LYS THR GLY ILE A
 ... CAAATATTACAACCAAAACAGGTGAATTA 2400
 ... 2380 2390

SN GLY GLU VAL LYS SER ALA SER GLY ASN ...
 ATGGCGAGTTAAATCAGCTTCGGTAATG...
 2410 2420 2430...
 ...VAL ASN ILE THR ALA SER GLY ASN THR LEU A
 ... TAAATATTACAGCGAGCGCAATACTTA 2460
 ... 2440 2450

69/235

FIG.20M

SN VAL SER ASN ILE THR GLY GLN ASN VAL ...
 A T G T A A G T A A C A T C A C T G G T C A A A T G T A A...
 2470 2480 2490...
 ...THR VAL THR ALA ASN SER GLY ALA ILE THR T
 ... C A G T A A C A G C A A A C T C A G G T G C C A T A C A A
 2500 2510 2520
 ...

HR THR GLU GLY SER THR ILE ASN ALA THR ...
 C C A C A G A A G G C T C A A C T A T T A C G C G A C A A...
 2530 2540 2550...
 ...THR GLY ASP ALA ASN ILE THR THR GLN THR G
 ... C A G G T G A T G C A A A T A T T A C A A C C C A A C A G
 2560 2570 2580
 ...

LY ASN ILE ASN GLY LYS VAL GLU SER SER ...
 G T A A T A T T A A T G G T A A A G T T G A A T C C A G T T...
 2590 2600 2610...
 ...SER GLY SER VAL THR LEU ILE ALA THR GLY G
 ... C T G G T T C T G T G A C G C T T A T T G C A A C T G G A C
 2620 2630 2640
 ...

LN THR LEU ALA VAL GLY ASN ILE SER GLY ...
 A A A C T C T T G C T G T A G G T A A T A T T T C A G G T G...
 2650 2660 2670...

70/235

FIG.20N

...ASP THR VAL THR ILE THR ALA ASP LYS GLY L
 ... A C A C T G T T A C C A T T A C T G C G G A T A A G G T A
 ... 2680 2690 2700

YS LEU THR THR GLN THR SER SER LYS ILE ...
 A A T A C C A C A C A A C A A G C T C T A A G A T T A ...
 2710 2720 2730...

...ASN GLY THR LYS SER VAL THR THR SER SER G
 ... A C G G A A C T A A G A G T G T A C C C A C C T C A A G C C
 ... 2740 2750 2760

LN SER GLY ASP ILE SER GLY THR ILE SER ...
 A A T C A G G T G A T A T T A G T G G C A C A T T C T G ...
 2770 2780 2790...

...GLY ASN THR VAL SER VAL SER ALA THR GLY S
 ... G T A A T A C G G T A A G C G T T A G T G C G A C C G G T A
 ... 2800 2810 2820

ER LEU THR THR GLN ALA GLY SER LYS ILE ...
 G C T T G A C C A C T C A A G C A G G C T C A A A A T T G ...
 2830 2840 2850...

...GLU ALA LYS THR GLY GLU ALA ASN VAL THR S
 ... A A G C A A A A C A G G T G A G G C T A A T G T A A C A A
 ... 2860 2870 2880

71/235

FIG.200

```

ER  ALA  THR  GLY  THR  ILE  GLY  THR  ILE  ...
    GCGCAACAGGTACAATTGGCGGTACAATCT...
    2890
    ...SER  GLY  ASN  THR  VAL  ASN  VAL  THR  ALA  ASN  T
    ...CTGGCAATACAGTAATAATTACAGCAATA
    ...
    2920
    2930
    2940

HR  ASP  ASN  LEU  THR  ILE  LYS  ASP  GLY  ALA  ...
    CTGATAATTAACTATTAAAGATGGCGCAA...
    2950
    ...ARG  ILE  LYS  ALA  THR  GLY  GLY  ALA  VAL  THR  L
    ...GAATTAAAGCAACGGGCGGAGCTGTGACTT
    ...
    2980
    2990
    3000

EU  THR  ALA  THR  GLY  THR  LEU  THR  THR  ...
    TAACCGCAACAGGAGGTACTTTAACCCAG...
    3010
    ...GLU  THR  SER  ASP  ILE  THR  SER  SER  ASN  G
    ...AAACAAGTTCTGTATATTATCCACAGCAATG
    ...
    3040
    3050
    3060

LY  GLN  THR  THR  LEU  THR  ALA  LYS  ASP  SER  ...
    GTCAGACAACCTCTCACGGCCAAAGGATAGCA...
    3070
    3080
    3090
    3100

```

72/235

FIG.20P

...SER ILE ALA GLY SER ILE ASN ALA ALA ASN V
 ... G T A T C G C A G G A A G C A T C A A T G C C G C C A A T G
 ... 3100 3110 3120

AL THR LEU ASN THR THR GLY THR LEU THR ...
 T G A C A T T A A A T A C C A C A G G C A C T T T A A C T A ...
 3130 3140 3150...

...THR VAL ALA GLY SER LYS ILE GLU ALA ALA S
 ... C T G T G G C A G G T T C A A A A A T C G A G G C A G C C A
 ... 3160 3170 3180

ER GLY THR LEU VAL ILE ASN ALA LYS ASP ...
 G T G G C A C C C T G G T T A T T A A T G C A A A A G A T G ...
 3190 3200 3210...

...ALA GLN LEU ASP GLY ALA ALA SER GLY ASP H
 ... C T C A G T T G G A C G G C G G C A T C A G G T G A C C
 ... 3220 3230 3240

IS THR VAL VAL ASN ALA THR ASN ALA ASN ...
 A C A C A G T A G T A A A T G C A A C C A A C G C A A C G ...
 3250 3260 3270...

...GLY SER GLY SER VAL ILE ALA THR THR SER S
 ... G C T C C G G C A G C G T A A T C G C G A C A C C T C A A
 ... 3280 3290 3300

73/235

FIG.20Q

```

ER ARG VAL ASN ILE THR GLY ASP LEU ILE ...
G C A G A G T G A A C A T C A C T G G G G A T T A A T C A A ...
3310
...THR ILE ASN GLY LEU ASN ILE ILE SER LYS A
... C A A T A A A T G G A T T A A A T A T C A T T T C A A A A
3320
... 3340 3350 3360

SN GLY LYS ASN THR VAL LEU LEU LYS GLY ...
A C G G T A A A A C A C C G T G C T G T T A A A G G T G ...
3370
...VAL GLU ILE ASP VAL LYS TYR ILE GLN PRO G
... T T G A A A T T G A T G T G A A A T A C A T T C A A C C G G
3380
... 3400 3410 3420

LY ILE ALA SER VAL ASN GLU VAL ILE GLU ...
G C A T A G C G A G C G T A A A T G A A G T A A T T G A A G ...
3430
...ALA LYS ARG ALA LEU GLU LYS VAL LYS ASP L
... C G A A A C G C G C C C T T G A G A A A G T A A A G A T T
3440
... 3460 3470 3480

EU SER ASP GLU GLU ARG GLU THR LEU ALA ...
T A T C T G A C G A A G A A G A G A A C A T T A G C T A ...
3490
... 3500 3510...
```

74/235

FIG.20R

...LYS LEU GLY VAL SER ALA VAL ARG PHE ALA G
 ... AACTTGCGGTGAGCGCTGTACGTTTGTCTG
 ... 3520 3530 3540

LU PRO ASN ALA ILE THR ILE ASN THR ...
 AGCCAAATAATGCCATTACGATTATAACAC...
 ... 3550 3560 3570...
 ...GLN ASN GLU PHE THR THR ARG PRO LEU SER G
 ... A A A T G A G T T T A C A C C A G A C C A T T A G T C
 ... 3580 3590 3600

IN VAL THR ILE SER GLU GLY LYS VAL CYS ...
 AAGTGACAAATTCTGTGAAGGTATGTT...
 ... 3610 3620 3630...
 ...PHE LEU ILE GLY ASN GLY ALA THR ILE CYS T
 ... TCTTAATCGGCAATGGCGCAACAATAATGCA
 ... 3640 3650 3660

HR ASN ILE ALA ASP ILE GLU ARG ***
 CCAATAATGCTGATATTGAGCGGTAG
 ... 3670 3680

75/235

FIG.21A

K21 *hmmw1A* sequence

```

LYS  GLU  TRP  LEU  LEU  ASP  PRO  ASP  ASP  ILE  ...
A A G A G T G G T T G T T A G A C C C G G A T G A T A T A ...
10                                     20      30...
...  ASN  ILE  VAL  ASN  GLY  SER  ASN  ILE  ASP  ALA
...  A A T A T T G T C A A C G G A A G T A A T A T T G A T G C T
...                                     40      50      60

GLN  LEU  GLN  PRO  GLY  ARG  GLY  ASP  THR  PRO  ...
C A A T T A C A G C C A G G T A G A G G C G A T A C A C C C ...
70                                     80      90...
...  ASN  LYS  VAL  SER  ALA  GLU  GLY  LEU  THR  SER
...  A A C A A G G T T T C A G C A G A A G G C T T A C A T C C
...                                     100     110     120

ILE  ASN  ASN  ALA  THR  LEU  SER  THR  ALA  LEU  ...
A T T A C A A T G C C A C A T T A T C C A C C G C T T T A ...
130                                     140     150...
...  GLN  LYS  GLY  ILE  GLU  VAL  ASN  ILE  SER  ALA
...  C A A A A G G G T A T T G A G G T C A A C A T T C T G C C
...                                     160     170     180

```

76/235

FIG.21B

THR LYS ASN VAL THR VAL ASN ALA ASP VAL ...
 A C A A A A A T G T A A C C G T C A A C G C G G A T G T T ...
 190 200 210...
 ...ASP VAL LYS ASN GLY THR LEU VAL LEU HIS
 ... G A T G T T A A A A C G G A A C A T T A G T A T T A C A T
 ... 220 230 240

SER GLN ARG ASN GLY VAL LYS ILE ASN GLY ...
 T C A C A A G G A A T G G A G T T A A A T T A A C G G T ...
 250 260 270...
 ... ASN ILE THR SER THR GLN ASN GLY ASN LEU
 ... A A T A T T A C C T C A C A C A A A A T G G T A A T T A
 ... 280 290 300

THR ILE LYS THR GLY GLY LYS VAL ASP VAL ...
 A C C A T T A A A C A G G T G G C A A G G T T G A T G T T ...
 310 320 330...
 ... HIS LYS ASN ILE THR LEU GLY MET GLY PHE
 ... C A T A A A A T A T C A C A C T T G G T A T G G G T T T
 ... 340 350 360

LEU ASN ILE THR SER ASP ASN ASN ILE THR ...
 T T G A A T A T T A C T T C C G A T A A T A C A T C A C C ...
 370 380 390...

77/ 235

FIG.21C

... PHE GLU LYS GLY ASP ASN LEU THR ILE THR
 ... TTGAAAAGGTGATTAATCTAACCATTAAC
 ... 400 410 420

ALA GLN GLY ASN ILE ILE SER ASN GLN GLU ...
 GCCCAGGAAATAATCTCTAATCAAGAG...
 ... 430 440 450...

... ASN LYS GLN LEU ARG PHE SER ASN VAL SER
 ... AATAACAACCTTAGATTAGTAATGTATCT
 ... 460 470 480

LEU ASN GLY MET GLY ALA GLY LEU THR PHE ...
 TTAATGGGATGGGTGCGGGTTTAACTTT...
 ... 490 500 510...

... THR ALA ASN LYS GLY ASN HIS THR HIS LYS
 ... ACTGCAATAAAGGTATAATCAATACCCATAAG
 ... 520 530 540

PHE ASP GLY THR LEU ASN ILE SER GLY LYS ...
 TTGATGGCACGCTTAACATTTCCGGAAAG...
 ... 550 560 570...

... VAL VAL ILE ASN GLN THR THR PRO HIS ASN
 ... GTAGTAATTAATCAACCAACCTCACAC
 ... 580 590 600

78/235

FIG.21D

```

ILE  ALA  PRO  TRP  ASN  ALA  SER  ALA  ASP  SER  ...
A  T  G  C  T  C  C  A  T  G  G  A  A  T  G  C  A  A  G  T  G  C  A  G  A  C  T  C  T  ...
610                                     620                                     630...
...  TYR  TRP  ASN  VAL  THR  THR  THR  LEU  THR  LEU  GLY
...  T  A  C  T  G  G  A  A  T  G  T  A  C  T  A  C  T  C  T  T  A  C  T  T  A  G  G  T
640                                     650                                     660
...

ASN  ASN  ALA  GLN  PHE  THR  PHE  ILE  LYS  PHE  ...
A  A  T  A  A  T  G  C  G  C  A  A  T  T  T  A  C  C  T  T  T  A  T  T  A  A  T  T  T  ...
670                                     680
...  VAL  ASP  SER  ASN  ARG  SER  VAL  ALA  LEU  ASN
...  G  T  C  G  A  T  A  G  C  A  A  C  C  G  C  T  C  G  G  T  A  G  C  T  C  T  A  A  T
690                                     700                                     710                                     720
...

SER  GLY  SER  ARG  SER  PHE  ALA  GLY  VAL  LYS  ...
A  G  C  G  G  T  T  C  A  A  G  A  A  G  T  T  T  G  C  G  G  G  G  T  A  A  G  ...
730                                     740                                     750...
...  PHE  TYR  GLY  LYS  ASN  ASN  ASN  GLU  MET  LYS  PHE
...  T  T  C  T  A  C  G  G  C  A  A  G  A  A  T  A  A  T  G  A  A  A  T  G  A  A  T  T  T
760                                     770                                     780
...

ASN  ILE  GLY  ASP  ASN  ALA  ASN  VAL  GLU  PHE  ...
A  A  T  A  T  T  G  G  T  G  A  T  A  A  T  G  C  T  A  A  T  G  T  T  G  A  A  T  T  C  ...
790                                     800                                     810...

```

79/235

FIG.21E

... LYS LEU LYS SER ASN ASP ASN THR SER ASN
 ... AAGTTAAATCAATGATAATACAGCAAC
 ... 820 830 840

ASN LYS PRO LEU PRO ILE GIN PHE LEU SER ...
 AACAACTACCAATTTCAGTTTTCATCT...
 ... 850 860 870...
 ... ASN ILE SER ALA THR GLY ASN GLY THR VAL
 ... AATATCTCAGCCACTGGTAATGGCACTGTA
 ... 880 890 900

SER PHE ASP ILE HIS ALA ASN LEU SER ALA ...
 TCTTTGATATACATGCCCAACTTGTCAGCA...
 ... 910 920 930...
 ... ARG SER THR GLU LEU ASN MET SER LEU ILE
 ... AAGTCACTGAGTTAATAATGAGTTTAAATT
 ... 940 950 960

ASN ILE SER ASN GLY VAL ASN PHE SER ILE ...
 AACATTCCTAATGGGGTTAATTTTCCTATA...
 ... 970 980 990...
 ... ASN SER HIS VAL ARG GLY ASN ASN ALA PHE
 ... AACTCCCATGTTCCGGGTAAATAATGCTTTT
 ... 1000 1010 1020

80/235

FIG.21F

GLU ILE LYS LYS ASP LEU ILE ILE ASN ALA ...
 G A A T C A A A A A G A T T T A A T T A A A T G C A ...
 1030 1040 1050...
 ... THR GLY SER ASN PHE ASN LEU LYS GLN THR
 ... A C T G G C T C G A A T T T A A T C T T A A G C A A A C G
 ... 1060 1070 1080

LYS ASP LYS PHE ASP ASN SER TYR GLU LYS ...
 A A G A T A A A T T T G A C A A T A G T T A C G A A A A ...
 1090 1100 1110...
 ... ASN ALA ILE PHE SER THR HIS ASN LEU THR
 ... A A C G C C A T T T C T C A C T C A T A A C C T A A C C
 ... 1120 1130 1140

ILE LEU GLY GLY ASN VAL THR LEU GLY GLY ...
 A T T C T T G G C G G C A A T G T T A C T C T A G G T G G G ...
 1150 1160 1170...
 ... GLU ASN SER SER SER ASN ILE LYS GLY ASN
 ... G A A A T T C A A G T A G T A A T A T T A A G G A A A T
 ... 1180 1190 1200

ILE ASN ILE ASN SER LYS ALA ASN VAL THR ...
 A T C A A C A T C A A T A G C A A G G C A A A T G T T A C A ...
 1210 1220 1230...

81/235

FIG.21G

... LEU GIN ALA HIS ALA GLY THR SER HIS LEU
 ... TTACAAAGCTCATGCCGGCCACGAGTCACCTT
 ... 1240 1250 1260

ASP LYS LYS GLU ARG THR LEU THR LEU GLY ...
 GATAAAGAGAGAGACCCCTAACCCCTTGGC...
 1270 1280 1290...
 ... ASN VAL SER VAL GLY GLY ASN LEU ASN ILE
 ... AATGTATCTGTGTGGGGGAAATTAAACATA
 ... 1300 1310 1320

ILE GLY SER ASN ALA HIS ILE ASP GLY ASN ...
 ATGGGCTCAAAATGCACATATTGACGGCAAT...
 1330 1340 1350...
 ... LEU SER ILE ALA GLU SER ALA LYS PHE GLN
 ... CTTCTATTGCAGAAAGTGCTAAATTTCAA
 ... 1360 1370 1380

GLY LYS THR ASN ASN ALA ASN LEU ASN ILE THR ...
 GGAAACCAATAACAACCTAATAATTACC...
 1390 1400 1410...
 ... GLY THR PHE THR ASN ASN GLY THR ALA ASP
 ... GGCACTTTTACCACACACGGCACCGCCGAC
 ... 1420 1430 1440

82/235

FIG.21H

```

ILE ASN ILE LYS GLN GLY VAL LYS LEU ...
A T T A A T A A A C A G G A G T G G T A A A C T C ...
1450                                     1460
... GLN GLY ASP ILE THR ASN ASN GLY ASN LEU
... C A A G G T G A T A T T A C C A A T A A C G G T A A T T A
1480                                     1490
...                                     1500

```

```

ASN ILE THR THR ASN ALA SER VAL ASN GLN ...
A A T A T C A C T A C G C C T C A G T C A A T C A A ...
1510                                     1520
... LYS THR ILE ILE ASN GLY ASN ILE THR ASN
... A A A C C A T T A T T A C G G A A A T A T A C T A A C
1540                                     1550
...                                     1560

```

```

LYS LYS GLY ASP LEU ASN ILE LYS ASP ILE ...
A A A A A G G C G A C T T A A C A T C A G G A T A T ...
1570                                     1580
... LYS ALA ASN ALA GLU ILE GLN ILE GLY GLY
... A A A G C C A A C G C C G A A A T C C A A A T T G G C G G C
1600                                     1610
...                                     1620

```

```

ASN ILE SER GLN LYS GLU GLY ASN LEU THR ...
A A T A T C T C G C A A A A G A A G G T A A T C T C A C G ...
1630                                     1640
...                                     1650...

```

83/235

FIG.21I

```

... ILE SER SER ASP LYS ILE ASN ILE THR LYS
... A T T C T C T G A C A A A T T A A T A T C A C C A A
... 1660 1670 1680

...
... GLY ASN SER ASP SER GLY VAL ALA SER ASN
... G G G A A T T C T G A T T C A G G C G T A G C A A G T A A T
... 1720 1730 1740

...
... THR LEU THR ASP ASN LEU ASN ILE SER GLY
... A C A T T A C A G A C A A T C T A A A C A T T T C A G G T
... 1780 1790 1800

...
... ASN SER ASP LEU ILE ILE GLY LYS ALA SER
... A A C A G T G A T T T A A T T A T T G G C A A G G C T A G C
... 1840 1850 1860

...
... THR LEU THR ASP ...
... T A C A G C T A A A G A T ...
... 1820 1830 ...
... ASN SER ASP LEU ILE ILE GLY LYS ALA SER
... A A C A G T G A T T T A A T T A T T G G C A A G G C T A G C
... 1840 1850 1860

```

84/235

FIG.21J

SER ASP ASN SER ASN ALA LYS GLN ILE THR ...
 A G T G A C A A C A G T A A T G C T A A C A A A T A C C ...
 1870 1880 1890...
 ... PHE ASP LYS VAL LYS ASP SER LYS ILE SER
 ... T T T G A C A A G G T T A A G A T T C A A A A T C T C A
 ... 1900 1910 1920

ALA GLY ASN HIS ASN VAL THR LEU ASN SER ...
 G C T G G C A A T C A C A A T G T A A C A C T A A A T A G C ...
 1930 1940 1950...
 ... LYS VAL GLU THR SER ASN SER ASP GLY SER
 ... A A A G T G G A A A C G T C T A A T A G C G A T G G T A G C
 ... 1960 1970 1980

THR GLY ASN GLY SER ASP ASP ASN ASN ILE ...
 A C C G G A A A C G G T A G C G A T G A C A A C A A T A T C ...
 1990 2000 2010...
 ... GLY LEU THR ILE SER ALA LYS ASP VAL THR
 ... G G C T T A A C T A T T C C G C A A A A G A T G T A A C G
 ... 2020 2030 2040

VAL ASN SER ASN ILE THR SER HIS LYS THR ...
 G T A A T A G T A A T A T C A C C T C T C A C A A A C A ...
 2050 2060 2070...

85/235

FIG.21K

... VAL ASN ILE SER ALA SER GLU GLY GLY ILE
 ... GTAAATATCTCTGCAATCAGAGGATATC
 ... 2080 2090 2100

THR THR LYS ALA GLY THR THR ILE ASN ALA ...
 ACTACTAAGCAGGCACACCAATTAATGCG...
 2110 2120 2130...
 ... THR THR GLY SER VAL GLU VAL THR ALA LYS
 ... ACCACAGGTAGCGTGGAAAGTAAC TGCTAAA
 ... 2140 2150 2160

THR GLY ASP ILE SER GLY THR THR ILE SER GLY ...
 AAGGCGATATTAGCGGTACGATTTCCGGT...
 2170 2180 2190...
 ... LYS THR VAL SER VAL THR ALA THR THR ASP
 ... AAGACAGTAAGTGTTACAGCAACCAACGAC
 ... 2200 2210 2220

SER LEU THR VAL LYS GLY GLY ALA LYS ILE ...
 AGTTTAACTGTTAAGGTGGCGCAAAATT...
 2230 2240 2250...
 ... ASN ALA THR GLU GLY THR ALA THR LEU THR
 ... AATGCGACAGAGGAAC TGCAACCTTAACT
 ... 2260 2270 2280

86/235

FIG.21L

ALA SER SER GLY LYS LEU THR THR GLU ALA ...
 G C A T C A T C G G G C A A A T T A C C C A C C G A G G C C ...
 2290 2300 2310...
 ... ASN SER ALA ILE SER GLY ALA ASN GLY VAL
 ... A A C T C T G C G A T T A G C G G G C T A A C G G T G T A
 2320 2330 2340
 ...

THR ALA SER SER GLN SER GLY ASP ILE SER ...
 A C T G C C T C A A G T C A A T C A G G C G A T A T T A G C ...
 2350 2360 2370...
 ... GLY THR ILE SER GLY LYS THR VAL SER VAL
 ... G G T A C G A T T T C C G G T A A G A C A G T A A G T G T T
 2380 2390 2400
 ...

THR ALA SER SER GLY SER LEU THR VAL GLY ...
 A C A G C A A G C T C T G G C A G T T T A A C T G T T G G A ...
 2410 2420 2430...
 ... GLY ASP ALA LYS ILE ASN ALA THR GLU GLY
 ... G G T G A C G C A A A A T T A T G C G A C A G A G G A
 2440 2450 2460
 ...

ALA ALA THR LEU THR ALA THR LYS GLY THR ...
 G C T G C G A C T T T A A C T G C A A C A A A G G C A C T ...
 2470 2480 2490...

87/235

FIG.21M

```

... LEU THR THR VAL LYS GLY SER ASN ILE ASP
... TTA ACTACCGTGTAAGGTTCAACATTGAC
...
2500 2510 2520

ALA ASN GLU GLY THR LEU VAL ILE ASN ALA ...
GCAACGAGGCACTTAGTTATTACGCA...
2530 2540 2550...
... GIN ASP ALA THR LEU ASN GLY ASP ALA SER
... CAGACGCCACCTAATA TGGTGATGCTCA
...
2560 2570 2580

GLY ASP ARG THR GLU VAL ASN ALA VAL ASN ...
GGCGACCGTACAGAGTGAAATGCAGTCAAC...
2590 2600 2610...
... ALA SER GLY SER GLY ASN VAL THR ALA LYS
... GCAAGCGGCTCTGGTAACGTAACTGCCGAA
...
2620 2630 2640

THR SER SER SER VAL ASN ILE THR GLY ASP ...
ACCTCAGCAGTG TGAAATATCACTGGAGAT...
2650 2660 2670...
... LEU SER THR ILE ASN GLY LEU ASN ILE ILE
... TTAAGCACCAATAAATGGA TTAAATATCAT
...
2680 2690 2700

```

88/235

FIG.21N

```

SER  LYS  ASN  GLY  LYS  ASN  THR  VAL  VAL  LEU  ...
TCGAAAATGGTAATAAACACCGTAGTTTA...
2710
...  LYS  GLY  ALA  GLU  ILE  ASP  VAL  LYS  TYR  ILE
...  AAGGTGCTGAATAATGATGTGAAATATATT
2720
...
2740
2750
2760

```

```

GLN  PRO  GLY  VAL  ALA  SER  ALA  ASN  GLU  VAL  ...
CAACGAGGTAGCAAGTGCGAATGAGGT...
2770
...  ILE  GLU  ALA  LYS  ARG  ALA  LEU  GLU  LYS  VAL
...  ATTGAAGCGAAGCGTGCCCTTGAAAGTA
2800
...
2810
2820

```

```

LYS  ASP  LEU  SER  ASP  GLU  GLU  ARG  GLU  THR  ...
AAGATTATCTGATGAAGAAAGAGAAACA...
2830
...  LEU  ALA  LYS  LEU  GLY  VAL  SER  ALA  VAL  ARG
...  TTAGCTAAACTTGGTGTAAGTGTGTACGT
2860
...
2870
2880

```

```

PHE  ILE  GLU  PRO  ASN  ASN  THR  ILE  THR  VAL  ...
TTATTGAACCAATAATACCATTACGGTT...
2890
2900
2910...

```

89/235

FIG.210

... ASN THR GLN ASN GLU PHE THR THR ARG PRO
 ... AACACACAATAAGATTATCAACCAAGACCA
 ... 2920 2930 2940

SER SER GLN VAL THR ILE SER GLU GLY LYS ...
 TCAGTCAAGTGACAAATTTCGTGAAGGTAGA...
 2950 2960 2970...

... ALA CYS PHE SER SER GLY ASN GLY ALA ALA
 ... GCGTGTTTCTCAAGTGGTAATGGCGCAACA
 ... 2980 2990 3000

VAL CYS THR ASN VAL ALA ASP ASP GLY GLN ...
 GTATGTACCAATGTTGCTGACGATGGACAG...
 3010 3020 3030

... GLN ***
 ... CAGTAG

90/235

FIG.22A

LCDC2 *hmm1A* sequence

```

  LYS  GLU  TRP  LEU  LEU  ASP  PRO  ASP  GLU  ...
  A A G A G T G G C T A C T G G A C C C T G A T G A A G ...
      10
      ... VAL  THR  ILE  GLY  ALA  GLY  ASP  VAL  GLY  ARG  SER
      ... T A A C T A T T G G A G C A G G T G A C G T A G G A C G T A G C
      ... 30      40      50      60

  ASP  ASP  SER  SER  ASP  THR  ALA  PHE  PRO  ...
  G A T G A T T C A A G T G A C A C T G C T T T C C C T A ...
      70
      ... THR  GLY  THR  GLY  GLU  ARG  ASN  SER  PRO  LYS  THR
      ... C C G G A A C A G G G G A A G A A A C A G C C C A A A C A
      ... 90      100      110      120

  ASN  ALA  GLN  ASN  ARG  PRO  THR  ILE  THR  ...
  A A C G C T C A A A C A G A C C A A C A A T A A C A A ...
      130
      ... ASN  THR  SER  LEU  GLU  GLN  ILE  LEU  LYS  ASN  GLY
      ... A C A C A T C T C T T G A G C A A A T A T T A A A A A T G G C
      ... 150      160      170      180

```

91/235

FIG.22B

```

THR  PHE  VAL  ASN  ILE  THR  ALA  LYS  ASN  ...
ACC  TTG  TTA  ACA  TAA  CCG  CCA  AAA  ATA...
190                                     ...
...LYS  ILE  LEU  VAL  ASN  SER  ASP  ILE  ASN  ILE  LYS
...  AAT  CTT  AGT  TAA  TAG  CGA  CAT  CAA  TAT  CAA  A
... 210                                     220 230 240

GLU  ASN  SER  HIS  LEU  ILE  LEU  TRP  SER  ...
GAG  AAC  TCC  CAC  CTA  ATC  CTC  TGG  AGC  G...
250                                     ...
...GLU  ARG  ASP  GLY  ASN  SER  GLY  VAL  GLN  ILE  ASP
...  AAG  AGA  TGG  CAA  CAG  CGC  GTT  CAG  ATT  GAT
... 270                                     280 290 300

GLY  ASN  ILE  THR  SER  ALA  THR  GLY  GLY  ...
GGC  AAT  ATT  ACT  TCC  GCT  ACT  GCG  GAA  A...
310                                     ...
...SER  LEU  THR  VAL  TYR  SER  SER  GLY  TRP  VAL  ASP
...  GCT  TAA  CCG  TTT  ACT  CTA  GTG  GCT  GGT  TGA  T
... 330                                     340 350 360

VAL  HIS  LYS  ASN  ILE  THR  LEU  ASN  SER  ...
GTC  ATA  AAA  AAC  ATT  ACA  CTT  AAT  TCAG...
370                                     380
...

```

92/235

FIG.22C

...GLY TYR LEU ASN ILE THR THR LYS SER GLY ASP
 ... GGTA CTTAAACA TTA C GACTAA AAGTGGAGAT
 ... 390 400 410 420

VAL ALA PHE GLU GLN GLY ASN ASP LEU ...
 GTCGCC TTCGAACAAGGGAATGACCTAA...
 430 440

...THR ILE THR GLY GLN GLY THR ILE THR ALA SER
 ... CCATTACAGGTC AAGGACTATTACCGCAAGC
 ... 450 460 470 480

LYS LYS GLY PHE ARG PHE ASP ASN VAL ...
 AAAAAGGTTT TAGATT TGAATAATGTTA...
 490 500

...THR LEU SER GLY VAL LYS LYS GLY PHE LEU PHE
 ... CTC TAGTGAGTGAA AAGGGTTCTCTTTT
 ... 510 520 530 540

LYS TYR SER GLN THR ASN ASN LYS ...
 AATA CAGCCAAACCAATAATAAG...
 550 560

...ASP SER ASN PHE GLU ASN HIS PHE ARG GLY THR
 ... ATAGCAATTTCGAA AACCATTTAGAGGAACT
 ... 570 580 590 600

93/235

FIG.22D

LEU ASN ILE SER GLY LYS VAL ASP ILE ...
 T T A A A T A T T T C A G G G A A A G T A G A T A T C T ...
 610 620
 ... LEU MET GLN ALA ARG GLN GLU ASN TRP ASN ARG
 ... T A A T G C A A G C A A G G C A G G A G A A C T G G A A C C G C
 ... 630 640 650 660

ARG HIS SER GLY ARG SER HIS TRP ASN ...
 A G A C A C T C G G G A C G C T C C C A C T G G A A T G ...
 670 680
 ... VAL THR ARG LEU ASN VAL SER THR ASN SER TYR
 ... T A C C C G A T T G A A T G T T T C T A C A A T A G T T A T
 ... 690 700 710 720

LEU ASN ILE THR ILE ASP ASN SER GLY ...
 C T C A A C A T C A C T A T T G A T A A C A G T G G C A ...
 730 740
 ... SER ARG PRO SER PRO GLY ALA GLY PRO LEU TYR
 ... G C C G T C C A T C C C C T G G T G C C G C C C T C T A T A T
 ... 750 760 770 780

ARG ARG SER GLY LEU ASN GLY ILE SER ...
 A G A C G T T C G G G T T T A A A T G G C A T A T C G T ...
 790 800 ...

94/235

FIG.22E

```

...PHE ASN ASN ASP THR VAL PHE ASN VAL ALA SER
... T T A A C A A T G A C A C T G T T T T A A T G T T G C G T C A
... 810 820 830 840

GLY SER ALA VAL ASN PHE SER ILE LYS ...
G G T C G G C A G T T A A C T T T A G C A T C A A G C ...
850 860
...PRO PRO ILE VAL SER ASN VAL HIS ASP GLY ASN
... C A C C A A T A G T A A G C A A T G T A C A C G A C G G A A T
... 870 880 890 900

HIS THR LEU PHE ASN GLY ASN VAL SER ...
C A C A C A T T A T T C A A T G G G A A T G T T T C A G ...
910 920
...VAL LEU GLY GLY GLY ASP VAL ASN PHE HIS PHE
... T T T A G G G G A G G G G A T G T C A A C T T C A T T T
... 930 940 950 960

ASN ALA SER SER SER ASN HIS TRP THR ...
A A C G C C T C C T C C A G C A A C C A C T G G A C T C ...
970 980
...HIS GLY VAL ILE LYS SER GLN ASN PHE ASN
... A T G G C C G T G G T T A T A A G T C T C A A A C T T T A A T
... 990 1000 1010 1020

```

95/235

FIG.22F

```

ALA SER GLU GLY SER LEU ARG PHE ...
GCC TCAGAGG GTCAG CTTAAG ATTCA...
1030
...LYS SER GLU GLY SER THR ARG THR PHE THR
...AAG CGAAG GTTCACACG AACCG CTTTACA
1050
1060
1070
1080

ILE GLU SER ASP LEU THR LEU ASN ALA ...
ATAG AAGTG ATTTA ACTTTAAATGCCA...
1090
...THR GLY GLY ASN ILE SER LEU ASN GLN VAL ALA
...CTGGGGCAATA TATCATTG AACCAAGTTGCCA
1110
1120
1130
1140

GLY ILE ASP GLY ASN LEU GLN LYS SER ...
GGTATTGATGG TAAATCTCTCCAA AAGCC...
1150
...LEU VAL ALA ASN LYS ASN ILE THR PHE GLU GLY
...TTGTAGCCAAATAAACA TAACTTTGAAGGG
1170
1180
1190
1200

GLY ASN ILE THR LEU ALA ALA ASP LYS ...
GGCAATAATCA CCGCTTGCAGCCGATAA A...
1210
1220

```

FIG.22G

...LYS PRO ILE GLU ILE LYS GLY ASN ILE THR VAL
 ...AACCAATAGAAATCAAGGTAAATTA CTGTT
 ...1230 1240 1250 1260

 LYS GLU GLY ALA ASN VAL THR LEU ARG ...
 AAGAGGAGCCAAATGTCACCTTCGTA...
 1270
 ...SER ALA ASN TYR GLY ASN ASP LYS SER ALA LEU
 ...GGCGAATTA TGGTAAATGACAAATCAGCTTA
 ...1290 1300 1310 1320

 SER ILE ARG GLY ASN VAL THR ASN LYS ...
 AGTATAGAGGAATAATGTCACCTAAAG...
 1330 1340
 ...GLY ASN LEU THR VAL THR GLY SER ALA ILE ASN
 ...GCAATCTCACCGTTACCGGCTCCGCTATCAAT
 ...1350 1360 1370 1380

 ILE GLU LYS ASN LEU THR VAL GLU GLY ...
 ATAGAAATAATCTTACCGTTTGAGGTA...
 1390 1400
 ...SER ALA LYS PHE LEU ALA ASN PRO ASN TYR SER
 ...GTGCTAAGTTT TAGCTAATCCAAATTA CAGC
 ...1410 1420 1430 1440

96/235

97/235

FIG.22H

PHE ASN VAL SER GLY LEU PHE ASP ASN ...
 TTTAACGTATCCGGCTATTTTGACAAACC...
 1450
 ...GLN GLY LYS SER ASN ILE SER ILE ALA LYS GLY
 ...AAGGCAAGTCAACATTTCCATTGCCCAAAGGA
 1470 1480 1490 1500

GLY ALA HIS PHE LYS ASP ILE ASN ASN ...
 GGGGCTCATTTAAGACATTATAACA...
 1510
 ...THR LYS SER LEU ASN ILE THR THR ASN SER ASP
 ...CTAAGAGTTTAAACATTACTACCACTCCGAC
 1530 1540 1550 1560

SER ALA TYR ARG THR ILE ILE GLU GLY ...
 TCCGCTTACCGCACCTATTATAGAGGCA...
 1570
 ...ASN ILE THR ASN SER ASN GLY ASP LEU ASN ILE
 ...ATAACCAACAGTACGGGGATTTAATAATC
 1590 1600 1610 1620

THR ASP ASN LYS ASN ASN ALA GLU ILE ...
 ACTGATAATAAATAACGCTGAATCC...
 1630 1640

98/235

FIG.22I

...GLN ILE GLY GLY ASN ILE SER GLN LYS GLU GLY
 ... A A A T T G G C G G C A A T A T C T C G C A A A A G A A G G T
 ...1650 1660 1670 1680

ASN LEU THR ILE SER SER ASP LYS ILE ...
 A A T C T C A C G A T T T C T T C C G A T A A A T T A ...
 1690 1700

...ASN ILE THR ASN GLN ILE THR ILE LYS LYS GLY
 ... A T A T C A C T A A C C A G A T A C A A T C A A G A A G G T
 ...1710 1720 1730 1740

VAL ASN LYS GLU ASP SER ASP SER SER ...
 G T T A A T A A G A G G A T T C T G A T T C A A G C A ...
 1750 1760

...THR ALA ASN ASN ALA ASN LEU THR ILE LYS THR
 ... C G G C A A A C A A T G C T A A T C T A A C C A T T A A A C C
 ...1770 1780 1790 1800

LYS GLU LEU LEU GLN LEU THR GLY ASP LEU ...
 A A G A A T T G C A A T T A A C G G G A G A C C T A A ...
 1810 1820

...ASN ILE SER GLY PHE ASP LYS ALA GLU ILE THR
 ... A T A T T C A G G C T T C G A T A A A G C A G A A T C A C A
 ...1830 1840 1850 1860

99/235

FIG.22J

```

ALA  LYS  GLU  GLY  ALA  ASP  LEU  ILE  ILE  ...
GCCAAGAGGGTGCCGATTTAATCATCG...
1870
...GLY  ASN  SER  ASP  ASN  ASN  ASN  ALA  ASN  ALA
...GTAAATAGTGATAATAACAATAATGCTAATGCT
...1890
1900
1910
1920

LYS  LYS  VAL  THR  PHE  ASN  GLN  VAL  LYS  ...
AAAAGTAACCTTTTAACCAAGTTAAAG...
1930
1940
...ASP  SER  LYS  ILE  SER  ALA  ASP  SER  HIS  ASN  VAL
...ATTCGAAATACTCTGCTGACAGTCACAAATGTA
...1950
1960
1970
1980

THR  LEU  ASN  SER  LYS  VAL  GLU  THR  SER  ...
ACACTAACAGTAAGTAGAACCCTCTA...
1990
2000
...ASN  GLY  ASN  ASN  ASP  ALA  GLU  SER  ASN  ASN  GLY
...ATGGCAATAATGACGCTGAAGCAATAATGGC
...2010
2020
2030
2040

ASP  GLY  THR  SER  LEU  THR  ILE  ASN  ALA  ...
GATGGCACCAAGCTTAACCTAATTAATGCAAA...
2050
2060
...
```

FIG.22K

SUBSTITUTE SHEET (RULE 26)

...LYS ASN ILE THR VAL ASN ASN ILE THR SER
 ... A A A T A T A C A G T A A C A C A A T A T T A C T T C T
 ...2070 2080 2090 2100

HIS LYS THR VAL ASN ILE THR ALA SER ...
 C A C A A A C A G T A A A T A T C A C T G C G T C A G ...
 2110 2120 ...

...GLU ASN VAL THR THR LYS ALA GLY THR THR ILE
 ... A A A T G T T A C C A C C A A A G C G G C A C A C C A T T
 ...2130 2140 2150 2160

ASN ALA THR THR GLY SER VAL GLU VAL ...
 A A T G C A A C C A C A G G T A G C G T A G A A G T A A ...
 2170 2180 ...

...THR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL
 ... C A G C C A A A A C A G G T G A T A T T A A A G G T A A A G T T
 ...2190 2200 2210 2220

GLU SER THR SER GLY SER VAL THR LEU ...
 G A A T C C A C T T C C G G C T C T G T A A C A C T T A ...
 2230 2240 ...

...THR ALA THR GLY GLU ALA LEU ALA VAL SER ASN
 ... C T G C A A C C G G A G A A G C T C T T G C T G T A A G C A A C
 ...2250 2260 2270 2280

100/ 235

09/8,06709

101/235

FIG.22L

```

ILE SER GLY ASN THR VAL THR ILE THR ...
A T T C A G G C A A C A C T G T T A C C A T C A C T G ...
2290
...ALA ASN LYS GLY LYS LEU THR THR GLN ALA GLY
... C A A A T A A G G G T A A T T A C A C T C A A G C A G G C
...2310 2320 2330 2340

```

```

SER THR VAL SER ALA ILE ASN GLY VAL ...
T C T A C G G T T A G C G C G A T T A A C G G T G T A A ...
2350
...THR ALA SER SER GLN SER GLY ASP ILE SER GLY
... C T G C C T C A A G C C A A T C A G G C G A T A T T A G C G G T
...2370 2380 2390 2400

```

```

THR ILE SER GLY ASN THR VAL LYS VAL ...
A C G A T T T C C G G T A A C A C A G T A A A A G T T A ...
2410
...SER ALA ILE GLY ASP LEU THR THR LYS SER GLY
... G T G C G A T C G G T G A T T T G A C T A C T A A T C C G G C
...2430 2440 2450 2460

```

```

SER GLU ILE LYS ALA LYS THR GLY GLU ...
T C G G A A A T C A A G G C A A A A C A G G T G A G G ...
2470 2480

```

FIG.22M

102/235

...ALA ASN VAL THR SER ALA THR GLY THR ILE GLY
 ...CTAACGTGACACAGTGCGACACAGGTACAAATTGGT
 ...2490 2500 2510 2520

GLY THR ILE SER GLY ASN ALA VAL ASN ...
 GGTAACGATTCTGCTGCTAATGCAATAATG...

2530

2540

...VAL THR ALA ASN THR GLY ASP LEU THR VAL GLU
 ...TTACAGCAATACTGGCGATTATACTGTTGAA
 ...2550 2560 2570 2580

ASP ALA ALA LYS ILE ASP ALA THR GLY ...
 GATGCCGCAAAATAATTGATGCGACAGGAG...

2590

2600

...GLY ALA ALA THR LEU THR ALA THR SER GLY LYS
 ...GAGCCGCGACCCCTAATCTGCACATCGGGCAAA
 ...2610 2620 2630 2640

LEU THR THR LYS ALA SER SER ILE ...
 TTAACCACTAAGGCTAGTTCAAGCATTAA...

2650

2660

...THR SER ALA ASN ASN GLN VAL ASN LEU SER ALA
 ...CTTCAGCTAATAACCAAGGTAAACCTTTTCAGCT
 ...2670 2680 2690 2700

103/235

FIG.22N

```

LYS  ASP  GLY  SER  ILE  GLY  GLY  ASN  ILE  ...
AAGGATGGTAGCATTGGGGGAAATATCA...
2710
...
...ASN  ALA  ALA  ASN  VAL  THR  LEU  ASN  THR  THR  GLY
...ATGCTGCTAATGTAACTGAACTACTACAGGC
...2730
2740
2750
2760

ALA  LEU  THR  THR  VAL  LYS  GLY  SER  SER  ...
GCTCTAACTACCGTGTAAGGGTTCAAGCA...
2770
...
...ILE  ASN  ALA  ASN  SER  GLY  THR  LEU  VAL  ILE  ASN
...TTAACGCAACACAGCGGCACCTTGTTATTAAAC
...2790
2800
2810
2820

ALA  LYS  ASP  ALA  GLU  LEU  ASN  GLY  GLU  ...
GCAAAAGACGCTGAGCTAAATGGTGAGG...
2830
...
...ALA  SER  GLY  ASN  HIS  THR  VAL  VAL  ASN  ALA  THR
...CATCAGGTAAACCATACTAGTAGTGAAATGCCAACCC
...2850
2860
2870
2880

ASN  ALA  ASN  GLY  SER  GLY  SER  VAL  ILE  ...
ACGCAAATGGGCTCCGGCAGCGTAATCG...
2890
2900
...
```

FIG.220

WO 00/20609

09/806709

PCT/CA99/00938

104/235

...ALA THR THR SER SER ARG VAL ASN ILE THR GLY
... CGACAACTCAAGCAGAGTGAAACATCCTGGG
...2910 2920 2930 2940

ASP LEU ILE THR ILE ASN GLY LEU ASN ...
GATTATCACCAATAAATGGAATTAAATA...

2950 2960

...ILE ILE SER LYS ASN GLY ILE ASN THR VAL LEU
... TCATTTCAAATAAACGGTATAAACACCGTACTG
...2970 2980 2990 3000

LEU LYS GLY VAL LYS ILE ASP VAL LYS ...
TTAAAGGCGTTTAAATAATTGATGTGAATA...

3010 3020

...TYR ILE GLN PRO GLY ILE ALA SER VAL ASP GLU
... ACATTCAACCGGGTATAGCAAGCGTAGATAA
...3030 3040 3050 3060

VAL ILE GLU ALA LYS ARG ILE LEU GLU ...
GTAAATTGAAGCGAAACGCATCCTTGAGA...

3070 3080

...LYS VAL LYS ASP LEU SER ASP GLU ARG GLU
... AGGTAAAGATTTCCTGATGAGAGAGAA
...3090 3100 3110 3120

FIG.22P

ALA LEU ALA LYS LEU GLY VAL SER ALA ...
 GCGTTAGCTAAACTTGCGCGTAAGCGCTG...
 3130
 ...
 ...VAL ARG PHE ALA GLU PRO ASN ALA ILE THR
 ...TACGTTTGTGCTGAGCCCAAATAATGCCCATTCAG
 ...3150 3160 3170 3180

105/235

ILE ASN THR GLN ASN GLU PHE THR THR ...
 ATTAATACACAAATAAGAGTTTACACCA...
 3190
 ...
 ...ARG PRO SER SER GLN VAL THR ILE SER GLU GLY
 ...GACCATCAAGTCAAGTGACACATTTCTGAGAGGT
 ...3210 3220 3230 3240

LYS VAL CYS PHE LEU ILE GLY ASN GLY ...
 AAGGTATGTTTCTTAAATCGGCAATGGTG...
 3250
 ...
 ...ALA THR ILE CYS THR ASN ILE ALA ASP ILE GLU
 ...CAACAATATGCACCAATAATGCTGATATTGAG
 ...3270 3280 3290 3300

ARG ***
 CCGTAG

FIG.23A

LCD2 *hmr2A* sequence

LYS GLU TRP LEU LEU ASP PRO ASP ASP ...
 A A G A G T G G T T G T T A G A C C C G G A T G A T G ...
 10
 ... VAL SER ILE ASP ALA PRO SER ALA GLU ARG THR
 ... T A T C C A T T G A C G C A C C T T C G G C T G A A C G C A C T
 ... 30 40 50 60
 106 / 235
 ASP THR GLY GLU ASP VAL GLU TYR THR ...
 G A C A C T G G C G A A G A C G T G G A A T A C A C C G ...
 70 80
 ... GLY THR GLY ALA ASP ILE ASN HIS GLN LYS GLN
 ... G A A C A G G G C T G A T A T T A A C C A T C A A A A C A A
 ... 90 100 110 120
 ASN SER GLU THR LYS SER THR LEU THR ...
 A A C A G C G A A A C C A A G T C A A C A T T A A C A A ...
 130 140
 ... ASN THR THR LEU GLU MET LEU LYS ARG GLY
 ... A C A C A C T C T T G A G G G A T G T T A A A A G G G G G
 ... 150 160 170 180

107/235

FIG.23B

LEU PHE VAL ASN ILE THR ALA ARG ASN ...
 C T T T T G T T A A T A T C A C C G C C A G A A A T A ...
 190 200
 ... LYS ILE ARG VAL ASN SER THR ILE ASN ILE GLY
 ... A A A T C C G A G T T A A T A G C A C C A T C A A T A T C G G G
 ... 210 220 230 240

ASP SER GLY HIS LEU THR LEU TYR LYS ...
 G A T A G C G G C C A T T T A A C C C T T T A C A A A A ...
 250 260
 ... LYS ARG LYS ASN ARG SER ASP GLY ILE GLN ILE
 ... A A A G A A A A A T C G T A G C G A T G G T A T T C A A A T T
 ... 270 280 290 300

ASN LYS ASP ILE THR SER THR GLY GLY ...
 A C A A G G A T A T T A C T T C T A C A G C G G A A ...
 310 320
 ... SER LEU THR ILE ASN SER ASP ASP TRP VAL ASP
 ... G T T T A A C T A T T A A C T C C G A C G A C T G G G T T G A T
 ... 330 340 350 360

ILE HIS GLY ASN ILE THR LEU GLY GLU ...
 A T T C A T G G A A A T A T C A C G C T T G G T G A G G ...
 370 380 ...

108/235

FIG.23C

...GLY PHE LEU ASN ILE THR SER SER ASP SER VAL
 ... G C T T T T A A A T A T T A C C T C T A G T G A T T C C G T G
 ... 390 400 410 420

ALA PHE GLU GLY GLY ASN GLY ASN LYS ...
 G C T T C G A G G G T G G A A A C G G C A A T A A A G ...

430 440

...GLY ARG SER SER ALA SER ALA GLN ILE ILE ALA
 ... G A C G T A G C T C A G C A A G T G C T C A A A T T A T C G C G
 ... 450 460 470 480

GLN GLY THR ILE THR LEU THR GLY GLU ...
 C A G G G T A C T A T A C T C T T A C T G G A G A A ...

490 500

...ASN LYS THR PHE ARG LEU ASN ASN VAL SER LEU
 ... A T A A A C C T T T A G A C T C A C A C A T G T G T C T T A
 ... 510 520 530 540

ASN GLY THR GLY ASN GLY LEU SER ILE ...
 A A T G G G A C G G G T A A T G G T C T A A G T A T T A ...

550 560

...ILE SER THR ALA SER ASN LEU SER HIS ARG LEU
 ... T T T C A A C A G C A A G C A A T T T A T C T C A T A G A C T T
 ... 570 580 590 600

109/235

FIG.23D

ASP GLY GLU ILE ASN VAL SER GLY ASN ...
 GACGGTGAAATTAAATGTAATCTGGAAATG...
 610
 ... VAL THR ILE ASN GLN THR THR GLN ASN ILE
 ... TAACAATTAAATCAAAACCACGCAAAACATT
 ... 630 640 650 660

GLU TYR TRP LYS ALA SER SER ASP SER ...
 GAATACGTGGAAGGCTAGCAGCGATTCTT...
 670
 ... TYR TRP ASN VAL THR SER PHE ASN LEU ARG GLU
 ... ATGGAAATGTCACCTCTCTTTAAATTGAGAGAA
 ... 690 700 710 720

ASP SER LYS PHE THR PHE ILE LYS TYR ...
 GATTCAAAGTTTACCCTTTATCAATAACG...
 730
 ... VAL ASN SER ALA ARG ASN GLY ASP VAL ARG GLY
 ... TTAACCTCTGCCAGAAATGGTGTAAGAGGA
 ... 750 760 770 780

ARG SER PHE ALA GLY VAL ILE PHE ASN ...
 AGAAGTTTTCAGGTGTGATATTTAATG...
 790 800 ...

FIG.23E

...ALA LYS GLY LEU THR THR SER PHE ASN VAL LYS
 ... C T A A A G G T C T C A C T A C A G C T T T A A C G T C A A G
 ... 810 820 830 840

LYS GLY SER THR VAL ASP PHE LYS LEU ...
 A A A G G C T C G A C A G T T G A T T T A A A T T A A ...

850

860

...LYS PRO ASN SER GLY TYR ASN SER GLN LYS ARG
 ... A G C C A A A T T C A G G C T A T A T T C A C A A A A A G G
 ... 870 880 890 900

110/235

ILE PRO ILE GLN PHE GLN SER ASN ILE ...
 A T T C C A A T T C A A T T C C A A T C C A C A T C T ...

910

920

...SER VAL SER GLY GLY GLY ARG VAL ASN ILE ASN
 ... C G G T C T C A G G A G G A G G T A A C A T T A A C
 ... 930 940 950 960

THR LEU ALA ASN LEU THR GLY GLY ...
 A C G C T C G C C A A T C T T A C A G G C G G A G G A ...

970

980

...VAL GLU ILE ARG SER SER SER ILE ASN VAL SER
 ... T T G A G A T A A G A T C G A G T T C A A T T A A T G T T C T
 ... 990 1000 1010 1020

FIG.23F

```

ASP  GLY  SER  THR  LEU  SER  MET  THR  ALA  ...
GATGGCTCAACCCCTCTCTATGACAGCTC...
1030
...
...GLN  ALA  ARG  ASP  ARG  ASN  ALA  PHE  GLU  ILE  THR
...AGGCTCGCGACAGGAATGCCCTTGAAATTATCC
1050
...1060
1070
1080

```

111/235

```

LYS  ASP  LEU  VAL  ILE  ASN  ALA  SER  ASN  ...
AAGATTTAGTTATAAACGCAGCAATT...
1090
...
...SER  ASN  LEU  SER  ILE  ILE  GLN  GIN  ASN  ASP  GLY
...CAACCTATCTATTATACAGCAAAATGATGGA
1110
...1120
1130
1140

```

```

PHE  ASP  ASN  ASN  GLN  LYS  ALA  ASN  ALA  ...
TTGATAATAATCAAAAGGCAAAATGCCA...
1150
...
...ILE  ASN  SER  LYS  TYR  ASN  VAL  THR  ILE  GLN  GLY
...TAACTCAAAATAATACGTAACTATTCAAGGT
1170
...1180
1190
1200

```

```

GLY  ASN  VAL  THR  LEU  GLY  GLY  ASN  ...
GGTAATGTTACCCCTTGGCGGGCAAAATT...
1210
1220
...

```

FIG.23G

112/235

...SER SER SER THR ILE THR GLY SER VAL ASN ILE
 ...C A G C A G T A C A A T C A C A G G A G T G T T A A T A T T
 ...1230 1240 1250 1260

GLY ALA ASN ALA ASN VAL THR LEU GLN ...
 G G C G C T A A T G C A A A T G T T A C T T T G C A A G ...
 1270 1280 ...
 ...ALA HIS ASN GLY ASN ASP ARG ASN LYS LYS LEU
 ...C C C A C A A T G G C A A T G A T A G A A A T A A A A G C T A
 ...1290 1300 1310 1320

THR PHE GLY ASN VAL SER VAL GLU GLY ...
 A C C T T C G G T A A T G T A T C T G T T G A A G G A G ...
 1330 1340 ...
 ...GLU LEU ARG LEU VAL GLY ALA SER ALA ASN ILE
 ...A A T T A A G G C T A G T T G G C G C A A G T G C A A A C A T T
 ...1350 1360 1370 1380

ASN ASN ASN LEU SER VAL LYS SER GLY ...
 A C A C A A A T C T T A G T G T T A A G A G C G G T G ...
 1390 1400 ...
 ...ALA LYS PHE LYS ALA GLU THR ASN ASP ASN LEU
 ...C T A A A T T C A A A G C A G A A C A A A T G A C A A C C T A
 ...1410 1420 1430 1440

FIG.23H

ASN ILE THR GLY THR PHE THR ASN ASN ...
 AACATTACCGGCACCTTTACCAACACG...
 1450
 ...GLY THR SER ILE ILE ASP VAL LYS GLY ALA
 ...GCACCTCCATAATTGATGTAAATAAGGGCCG
 1470
 ...1480
 1490
 1500

113/235

ALA LYS LEU GLY ASN ILE THR ASN ASP ...
 GCAAACTAGGCAATAATTACCAATGATG...
 1510
 ...GLY ASN LEU ASN ILE THR THR ASN ALA LYS ASN
 ...GTAATTATAATCACTACTAATGCTAAAC
 1530
 ...1540
 1550
 1560

GLY GLN LYS SER VAL ILE ASN GLY ASN ...
 GGTCAAAAGCGTTATCAACGGAATAA...
 1570
 ...ILE THR ASN LYS GLY ALA LEU ASN ILE THR
 ...TAACTAACAATAAGGTGCTTTAATAATTACG
 1590
 ...1600
 1610
 1620

ASN ASN GLY ASN ASP THR GLU ILE GLN ...
 AATAATGGTAATGACACTGAATAATCCAA...
 1630
 ...1640
 ...

114/235

FIG.23I

...ILE GLY GLY ASN ILE SER GIN LYS GLU GLY ASN
 ... TTGGCGGCAATAATCTCTCGCAAAAGAGGTAAAT
 ...1650 1660 1670 1680

LEU THR ILE SER SER ASP LYS ILE ASN ...
 CTCACGATTCTTCTTGACAAATAATA...
 1690 1700 ...
 ...ILE THR LYS ARG ILE GLU ILE LYS ALA GLY THR
 ... TCACCAACCGGATAGAAATTAAGGCAGGTACT
 ...1710 1720 1730 1740

ASP GIN GLY ASN SER ASP SER GLY VAL ...
 GATCAGGGGAATTCTGATTTCAGGCGTAG...
 1750 1760 ...
 ...ALA SER ASN ALA ASN LEU THR ILE LYS THR LYS
 ... CAGTAATGCTAATCTAACCATTAACCAAA
 ...1770 1780 1790 1800

GLU LEU LYS LYS LEU THR GLU ASN LEU ASN ...
 GAATTGAATAACAGAAACCTAATAA...
 1810 1820 ...
 ...ILE SER GLY PHE ASP LYS ALA GLU ILE VAL ALA
 ... TTTCAGGTTTTCATTAAGCAGAAATTGTAGCC
 ...1830 1840 1850 1860

FIG.23J

LYS GLU ASN ASN ASN LEU ILE ILE GLY ...
 A A G A G A A T A C A A T T T A A T T A T T G G C A ...
 1870 1880
 ... ASN ASN ASN GLY ASP ASN ALA ASN ALA LYS THR
 ... A T A A T A T G G C G A C A A T G C T A A C G C C A A A C A
 ...1890 1900 1910 1920

115/235

VAL THR PHE ASN ASN VAL LYS ASP SER ...
 G T A A C T T T A A C A A T G T T A A G A T T C A A ...
 1930 1940
 ... LYS ILE SER ALA ASN GLY HIS ASN VAL THR LEU
 ... A A A T C T C T G C T A A C G G T C A C A A T G T G A C A C T A
 ...1950 1960 1970 1980

ASN SER LYS VAL GLU THR SER ASP GLY ...
 A A T A G C A A A G T G G A A A C A T C T G A T G G A A ...
 1990 2000
 ... ASN SER ASN THR GLU GLY ASN SER ASP ASN ASN
 ... A C A G T A A C A C T G A A G G T A A T A G T G A C A A T A A C
 ...2010 2020 2030 2040

ALA GLY LEU THR ILE ASP ALA LYS ASN ...
 G C C G G C T T A A C T A T C G A T G C A A A A A T G ...
 2050 2060 ...

09/806709

FIG.23K

116/235

...VAL THR VAL ASN ASN ASP ILE THR SER HIS LYS
 ...TAA CAGTAAACAACGATATCACTCTCTCAACA
 ...2070 2080 2090 2100

THR VAL ASN ILE THR ALA SER GLU ARG ...
 A C A G T A A A T A T C A C T G C G T C A G A A A G G A ...
 2110 2120

...ILE ASP THR LYS LYS ALA ASP THR THR ILE ASN ALA
 ...TTGATACCTAAAGCTGATACCAACCATTAATGCA
 ...2130 2140 2150 2160

THR THR GLY ASN VAL LYS LEU THR ALA ...
 A C C A C C G G C A A C G T G A A A C T A A C A G C T G ...
 2170 2180

...VAL THR SER ASP ILE GLN GLY GLY ILE LYS SER
 ...TAA CAGTGATATCCAGGTGGAATTAATCT
 ...2190 2200 2210 2220

ASN SER GLY ASP VAL ASN ILE THR THR ...
 A A T T C T G G T G A T G T A A A T A T C A C A C C A ...
 2230 2240

...SER THR GLY SER ILE ASN GLY LYS ILE GLU SER
 ...GCCACAGGTAGCATTAACGGTAATAATTGAATCC
 ...2250 2260 2270 2280

FIG.23L

LYS SER GLY SER VAL THR LEU THR ALA ...
 AAGTCTGGCTCTGTAACTTACCGCAA...
 2290
 ...
 ...THR GLU LYS THR LEU THR VAL GLY ASN VAL SER
 ...CCGAAAACTCTTACTGTAGGCAATGTTTCG
 ...2310 2320 2330 2340

117/235

GLY ASN THR VAL THR VAL THR ALA ASN ...
 GGCAACACCGTTACTGTACTGCAATA...
 2350
 ...
 ...ARG GLY ALA LEU THR THR LEU ALA GLY SER THR
 ...GAGGTGCATTAAACCACCTTTGGCAGGCTCTACG
 ...2370 2380 2390 2400

ILE ASN GLY THR ASN GLY VAL THR THR ...
 ATTACGGGACTAACGGTGTAACCT...
 2410
 ...
 ...SER SER GLN SER GLY GLU ILE GLY GLU VAL
 ...CAAGTCAATCAGGCGAGATTGGCGGTGAGGTT
 ...2430 2440 2450 2460

THR GLY LYS THR VAL SER VAL THR ALA ...
 ACTGGTAAGACAGTAAGTGTACAGCAA...
 2470 2480 ...

FIG.23M

118/235

...THR ALA GLY SER LEU THR VAL LYS GLY GLY ALA
 ...CTGCCGGCAGCTTAACTGTTAAGGTGGCGCA
 ...2490 2500 2510 2520

LYS ILE ASN ALA THR GLU GLY THR ALA ...
 AAAATAATGCGACAGAACTGCAAA...

2530 2540

...THR LEU THR ALA SER SER GLY LYS LEU THR THR
 ...CCTTAACCTGCATCATCGGGCAAAATTACCAACC
 ...2550 2560 2570 2580

GLU ALA SER SER ASN ILE THR SER ALA ...
 GAGGCTAGCTCAACATCATCTTCAGCCA...

2590 2600

...LYS GLY GLN VAL ASP LEU SER ALA GLN ASP GLY
 ...AAGGTCAGGTAGACCTTTCAGCTCAGGATGGT
 ...2610 2620 2630 2640

SER ILE ALA GLY GLN ILE SER ALA ALA ...
 AGCATTGCAAGGACAAATTAGTGCACTA...

2650 2660

...ASN VAL THR LEU ASN THR THR GLY THR LEU THR
 ...ATGTAACTGAAATACTACAGGCACTCTAAT
 ...2670 2680 2690 2700

119/235

FIG.23N

THR VAL GLU GLY SER SER ILE ASN ALA ...
 ACCGTAGAGGTTTCAGCATTAACGCA...
 2710
 ...
 ...ASN GLU GLY THR LEU VAL ILE ASN ALA ASN ASP
 ...ACGAAGGCACCTTGTTTATAACGCAACGAC
 2720
 ...2730 2740 2750 2760

ALA LYS LEU ASP GLY LYS ALA SER GLY ...
 GCCAAGTTAGATGGTAAGGCATCAGGTA...
 2770
 ...
 ...ASN ARG THR GLU VAL ASN ALA THR ASN ALA SER
 ...ACCGTACAGAGTAATA TGCAACTAACGCCAAGC
 2780 2800 2810 2820
 ...2790

GLY SER GLY SER VAL THR ALA LYS THR ...
 GGCTCTGGTAGCGTGACTGCGAATAACCT...
 2830
 ...
 ...SER SER SER VAL ASN ILE THR GLY ASP LEU ASN
 ...CAGCAGCGTGAAATATCACCGGGATTTAAC
 2840 2850 2860 2870 2880
 ...2890

THR ILE ASN GLY LEU ASN ILE ILE SER ...
 ACAATAATGGGTTTAAATAATCATTTCCGG...
 2890
 ...

FIG.230

...GLU ASN GLY ARG ASN THR VAL ARG LEU ARG GLY
 ... A A A A T G G T A G A A A C A C T G T G C G C T T A A G A G G C
 ...2910 2920 2930 2940

 LYS GLU ILE GLU VAL LYS TYR ILE GLN ...
 A A G G A A A T T G A G G T G A A A T A T A T C C A G C ...
 2950 2960 ...
 ...PRO GLY VAL ALA SER VAL GLU GLU VAL ILE GLU
 ... C A G G T G T A G C A A G T G T A G A A G A A G T A A T T G A A
 ...2970 2980 2990 3000

 ALA LYS ARG VAL LEU GLU LYS VAL LYS ...
 G C G A A A C G C G T C C T T G A G A A A G T G A A A G ...
 3010 3020 ...
 ...ASP LEU SER ASP GLU GLU ARG GLU THR LEU ALA
 ... A T T A T C T G A T G A A G A A A G A G A A C A T T A G C T
 ...3030 3040 3050 3060

 LYS LEU GLY VAL SER ALA VAL ARG PHE ...
 A A C T T G G T G T A A G T G C T G T A C G T T T A ...
 3070 3080 ...
 ...ILE GLU PRO ASN THR ILE THR VAL ASN THR
 ... T T G A C C A A A T A T A C C A T T A C G G T T A A C A C A
 ...3090 3100 3110 3120

FIG. 23P

GLN ASN GLU PHE THR THR ARG PRO SER ...
 C A A A T G A G T T A C A A C C A C C A T C A A...
 3130
 3140
 ...
 ...SER GLN VAL THR ILE SER GLU GLY LYS ALA CYS
 ... G T C A A G T G A C A A T T C T G A A G G T A A G G C G T G T
 ...3150 3160 3170 3180
 ...3150
 PHE SER SER GLY ASN GLY ALA ALA VAL ...
 T T C T C A A G T G G T A A T G G C G C A G C A G T A T...
 3190
 3200
 ...
 ...CYS THR ASN VAL ALA ASP ASP GLY GLN GLN ***
 ... G T A C C A A T G T T G C T G A C G A T G G A C A G C A G T A G
 ...3210 3220 3230 3240

FIG.24A

PMH1 *hmmw1A* sequence

LYS GLU TRP LEU LEU ASP PRO ASP ASN ...
 A A G A G T G G T T A G A C C C G G A T A A T G ...
 10 20
 ... VAL ASN ILE VAL LYS GLY THR GLU LEU GLN ASN
 ... T C A A T A T T G T T A A A G G A A C C C G A A T T A C A G A A T
 ... 30 40 50 60
 122/235
 ASP LEU VAL VAL ARG GLY ASP SER ILE ...
 G A T T G G T T G T T A G G G G C G A T A G T A T T G ...
 70 80
 ... GLU LYS LYS ASN ALA PRO THR LYS THR THR ILE
 ... A G A A A A G A A T G C C C C T A C C A G A C T A C A A T T
 ...

123/235

FIG.24B

```

ASN  LYS  VAL  ASN  VAL  THR  THR  ASP  ILE  ...
A A T A A G T A A A T G T T A C T A C A G A T A T T A ...
190
...
...ASN  VAL  TYR  ASN  GLY  ALA  LEU  THR  LEU  HIS  SER
... A T G T T A T A A T G G A G C A T T A A C G T T A C A C T C A
210
... 220
230
240

```

```

GLU  ARG  ASP  GLY  VAL  GLU  ILE  ASN  GLY  ...
G A C G A G A T G G A G T T G A A A T T A A C G G T A ...
250
...
...ASN  ILE  THR  SER  GLU  LYS  ASN  GLY  ASN  LEU  THR
... A T A T T A C C T C A G A A A A A A T G G T A A T T A A C C
270
... 280
290
300

```

```

ILE  LYS  ALA  GLY  SER  TRP  VAL  ASP  VAL  ...
A T T A A G C A G G T A G C T G G G T T G A T G T T C ...
310
...
...HIS  LYS  ASN  ILE  THR  LEU  GLY  GLU  GLY  PHE  LEU
... A T A A A A T A T C A C A C T T G G C G A G G G T T T T G
330
... 340
350
360

```

```

ASN  ILE  THR  SER  GLY  ASP  ILE  ALA  PHE  ...
A A T A T T A C T T C C G G T G A T A T C G C C T T C G ...
370
... 380
...

```

124/235

FIG.24C

...GLU LYS GLY ASN ASN LEU THR ILE THR ALA GLN
 ... A A A A G G T A A T A T C T A C C A T T A C C G C T C A A
 ... 390 400 410 420

GLY ASN ILE THR SER ASN LYS ASP GLY ...
 G G A A A T A T A C C T C T A A T A A G A C G G A A ...
 430 440 ...
 ...LYS GLN LEU ARG LEU ASN ASN VAL SER LEU ASN
 ... A A C A A C T T A G A C T T A A T A A T G T A T C T T T A A A T
 ... 450 460 470 480

GLY THR GLY ALA GLY LEU ASN PHE ILE ...
 G G A C A G G T G C A G G T T T A A A C T T T A T T G ...
 490 500 ...
 ...ALA ASN GLN ASN ASN PHE THR HIS ASN ILE SER
 ... C A A A T C A A A T A A T T T A C A C A C A C A T T A G T
 ... 510 520 530 540

GLY ALA ILE ASN ILE SER GLY VAL VAL ...
 G G C G C G A T T A A C A T T T C C G G A G T A G T A A ...
 550 560 ...
 ...THR ILE ASN GLN THR THR THR LYS LYS ASN ALA LYS
 ... C G A T T A A T C A A A C T A C G A A A A A A C G C T A A G
 ... 570 580 590 600

125/235

FIG.24D

ALA TRP ASN THR SER TYR ASP SER TYR ...
 G C A T G G A A T A C A G C T A T G A C T C T T A C T ...
 ... 610
 ... TRP ASN VAL SER THR LEU THR LEU SER ASN ASP
 ... G G A A C G T A T C T A C T C T T A C T T T A G C A A T G A T
 ... 630 650 660

ALA LYS PHE THR PHE ILE LYS TYR VAL ...
 G C G A A A T T A C C T T T A T T A A A T A T G T C G ...
 ... 670
 ... ASP SER ASN HIS SER THR ASN SER ASP SER
 ... A C A G C A A T C A T T C G A C A A A C T C C A G T G A T T C A
 ... 690 700 710 720

ARG SER PHE ALA GLY VAL LYS PHE HIS ...
 C G A A G T T T T G C G G A G T A A A G T T C C A C G ...
 ... 730
 ... GLY LYS ASN ASN GLU MET LYS PHE ASN ILE GLY
 ... G C A A G A A T A A T G A A A A T G A A A T T T A A T A T T G G T
 ... 750 760 770 780

ASN ASN ALA LYS ALA GLU PHE ARG LEU ...
 A A T A A T G C C A A G G C T G A A T T T A G G T T A A ...
 ... 790 800 ...

FIG.24E

126/235

...LYS PRO ASN GLU LYS THR THR PRO ASN ARG PRO
 ... A A C C A A A T G A G A G A C A C T C C T A A C A G A C C A
 ... 810 820 830 840

LEU PRO ILE GLN PHE LEU SER ASN ILE ...
 C T A C C A A T T C A G T T T T A T C T A A T A T T T ...

850

860

...SER VAL THR GLY GLY GLY SER VAL PHE ASP
 ... C G G T C A C T G G C G G A G G T T C T G T G T T T T C G A T
 ... 870 880 890 900

ILE TYR ALA ASN LEU TRP GLY LYS GLY ...
 A T A T A C G C T A A C C T T T G G G G T A A A G G G A ...

910

920

...THR GLU LEU LYS MET ASP SER ILE ASN VAL SER
 ... C T G A G C T A A A G A T G G A T T C A A T T A A C G T T C T
 ... 930 940 950 960

SER GLY SER ASN LEU THR LEU SER ...
 A G C G G C T C T A A T C T T A C C T T A A A T T C C C ...

970

980

...HIS VAL ARG LYS TYR ASN ALA PHE GLU ILE ASN
 ... A T G T T C G C A A G T A T A A T G C T T T T G A A A T C A A T
 ... 990 1000 1010 1020

127/235

FIG.24F

LYS ASP LEU THR ILE ASN ALA THR ASN ...
 A A G A C T T A A C T A T A A C G C A A C T A A T T ...
 1030
 ...
 ...SER ASN PHE ASN LEU ARG GLN THR SER ASP SER
 ... C A A A T T C A A C C T C A G A C A G A C G T C A G A T A G T
 1050 1060 1070 1080

PHE ARG ASN GLY TYR ARG ASN ASN ALA ...
 T T T C G T A A C G G G T A C C G C A A T A A T G C C A ...
 1090
 ...
 ...ILE ASN SER THR HIS ASN ILE SER ILE LEU GLY
 ... T C A A T T C A A C C C A C A C A T A T C C A T C T T G G C
 1110 1120 1130 1140

GLY ASN VAL THR LEU GLY GLN ASN ...
 G G C A A C G T C A C T C T C G G C G G A C A A A C T ...
 1150
 ...
 ...SER SER SER ILE MET GLY ASN ILE ILE ILE
 ... C A G C A G C A G C A T T A T G G G A A T A T C A T C A T C
 1170 1180 1190 1200

LYS ARG ALA ALA ASN VAL THR LEU GLU ...
 A A G C G A G C A G C A A A T G T T A C G C T A G A G ...
 1210 1220

FIG.24G

...ALA ASP ASN SER HIS ASN SER ASP ASN VAL LYS
 ... CCGATAATAGTCACAAATCTCTGACACACGTAAAG
 ...1230 1240 1250 1260

ASP ARG THR ILE ASN LEU GLY ASN LEU ...
 GATAGAACTATAATACTTTGGCAACTTGA...

1270

1280

...THR VAL GLU GLY ASN LEU SER LEU ILE GLY GLU
 ... CCGTTGAGGGGAAATTTAAGTTTAAATTGGCGGAA
 ...1290 1300 1310 1320

128/235

ASN ALA ASN ILE ASN GLY ASN LEU SER ...
 AATGCCAAATAATTACGGGCAATCTCTCCA...

1330

1340

...ILE GLU LYS GLU ALA ILE PHE LYS GLY LYS THR
 ... TTGAAAGAGAGCCATCTTTAAAGGAAACCC
 ...1350 1360 1370 1380

LYS ASP SER LEU ASN ILE THR GLY ASN ...
 AAGGACAGCCCTAAACATCACCGGCAACT...

1390

1400

...PHE THR ASN ASN GLY THR ALA GLU ILE ASN ILE
 ... TTACCAATAATGGCACCTGCCGAAATTAATA
 ...1410 1420 1430 1440

129/235

FIG.24H

SER GIN GLY VAL VAL SER LEU GLY ASP ...
 AGCCAAAGGAGTGGTAAGTCTTGCGGATA...
 1450
 ...
 ...ILE THR ASN ASP GLY LYS LEU ASN ILE THR THR
 ...TTACCAATGATGGCAAAATTAAACATCACCACT
 1480
 ...1470 1490 1500

HIS ALA LYS SER GLY GIN LYS SER ILE ...
 CACGCCAAGAGCGGTCAAAAGCATTA...
 1510
 ...
 ...ILE ARG GLY ASP ILE ILE ASN LYS GIN GLY ASN
 ...TCCGCGGAGATATAATTAAACAAGGGAAT
 1540 1550 1560
 ...1530

LEU ASN ILE THR ASP ASN ASN SER ASN ...
 TTAATAATTACGGACAAATAATAGTAATG...
 1570
 ...
 ...ALA GLU ILE GLU ILE GLY GLY ASN ILE SER GIN
 ...CTGAATTGAATAATGGCGGCAATATCTCGCAA
 1580 1590 1600 1610 1620
 ...1570

LYS GLU GLY ASN LEU THR ILE SER SER ...
 AAGAGGTAATCTCACCAATTTCTCTCTG...
 1630
 ... 1640

FIG.24I

SUBSTITUTE SHEET (RULE 26)

130/235

...ASP LYS VAL ASN ILE THR LYS GLN ILE THR ILE
 ... A T A A G T C A A T A T T A C C A A A C A G A T A A C A A T C
 ...1650 1660 1670 1680

LYS ALA GLY VAL ASP GLY GLU SER SER ...
 A A G C A G G C G T T G A T G G G A G A G T T C T A...

1690

1700

...SER SER SER THR ALA SER ASP ALA ASN LEU THR
 ... G T T C A A G C A C A G C A A G T G A T G C C A A T C T A A C C
 ...1710 1720 1730 1740

ILE LYS THR LYS GLU LEU THR LEU THR ...
 A T T A A A C C A A A G A G T T A A C A T T A A C A G...

1750

1760

...ASP ASN LEU ASN ILE SER GLY PHE ASN LYS ALA
 ... A C A A T C T A A A C A T T T C A G G T T T A A T A A A G C A
 ...1770 1780 1790 1800

GLU ILE THR ALA LYS ASP ASN SER ASP ...
 G A A T T A C A G C T A A A G A T A A C A G T G A T...

1810

1820

...LEU ILE ILE GLY LYS ALA SER SER ASP ASN SER
 ... T A A T T A T T G G C A A G G C T A G C A G T G A C A C A G T
 ...1830 1840 1850 1860

131/235

FIG.24J

ASN ALA LYS GLN VAL THR PHE ASP LYS ...
 AATGCTAAACAAGTAACCTTTTGACCAAGG...
 1870 1880
 ... VAL LYS ASP SER LYS ILE SER ALA GLY ASN HIS
 ...TTAAAGATTCAAAATACTCAGCTGGCAATCAC
 ...1890 1900 1910 1920

ASN VAL THR LEU ASN SER LYS VAL GLU ...
 AATGTAAACACTAAATAAGCAAGTGGAATA...
 1930 1940
 ... THR SER ASN SER ASP GLY SER THR GLY ASN GLY
 ...CGTCTAATAAGCGATGGTAGCACCGGAACGGT
 ...1950 1960 1970 1980

SER ASP ASP ASN ASN ILE GLY LEU THR ...
 AGCGATGACACAATAATCGGCTTAACTA...
 1990 2000
 ... ILE SER ALA LYS ASP VAL THR VAL ASN SER ASN
 ...TTTCGGCAAAAGATGTAAACGGTAATAGTAAT
 ...2010 2020 2030 2040

ILE THR SER HIS LYS THR VAL ASN ILE ...
 ATCACTCTCACAAACAGTAATATCT...
 2050 2060 ...

FIG.24K

WO 00/20609

09/806709

PCT/CA99/00938

...SER ALA SER GLU GLY GLY ILE THR THR LYS ALA
... CTGCA TCAGAGGATCTACTAAGCA
...2070 2080 2090 2100

GLY THR THR ILE ASN ALA THR THR GLY ...
GGCAACCAATTAA TGCGACCAAGGTA...

2110

2120

...SER VAL GLU VAL THR ALA LYS THR GLY ASP ILE
... GCGTGGAAGTAAC TGCTAAACAGCGATT
...2130 2140 2150 2160

132/235

SER GLY THR ILE SER GLY LYS THR VAL ...
AGCGGTACGATTTC CGGTAGACAGTA...

2170

2180

...SER VAL THR ALA SER THR GLY ASP LEU THR VAL
... GTGTTACAGCAAGCACTGGCGATTTAACGT
...2190 2200 2210 2220

ARG LYS ALA ALA THR ILE SER ALA THR ...
AGGAAGCTGCACCAATTAGTGCGACAG...

2230

2240

...GLU GLY ALA ALA THR LEU THR ALA THR GLY ASN
... AAGGAGCTGCACCACTTAACCGCAACAGGGAAT
...2250 2260 2270 2280

FIG.24L

THR LEU THR THR GLU ALA GLY SER SER ...
 A C C T T G A C T A C T G A A G C C G G T T C T A G C A ...
 2290 2300 ...
 ...ILE THR SER THR LYS GLY GLN VAL ASP LEU SER
 ... T C A C T T C A A C T A A G G G T C A G G T A G A C C T T T C A
 ...2310 2320 2330 2340

133/235

ALA GLN ASP GLY SER ILE ALA GLY GLN ...
 G C T C A G G A T G G T A G C A T T G C A G G A C A A ...
 2350 2360 ...
 ...ILE SER ALA ALA ASN VAL THR LEU ASN THR THR
 ... T T A G T G C A G C T A A T G T G A C A T T A A T A C C A C A
 ...2370 2380 2390 2400

GLY THR LEU THR THR VAL GLU GLY SER ...
 G G C A C C T T A A C T A C T G T A G A A G G T T C A A ...
 2410 2420 ...
 ...ASN ILE LYS ALA THR SER GLY THR LEU ALA ILE
 ... A C A T T A A G G C A A C C A G T G G C A C C T T A G C T A T T
 ...2430 2440 2450 2460

ASN ALA LYS ASP ALA LYS LEU ASP GLY ...
 A A C G C A A A A G A C G C T A A G C T A G A T G G T A ...
 2470 2480 ...

FIG.24M

134/235

...THR ALA SER GLY ASN ARG THR GLU VAL ASN ALA
 ... C G G C A T C A G G T A A C C G T A C A G A A G T A A A T G C A
 ...2490 2500 2510 2520

THR ASN ALA SER GLY SER VAL ...
 A C T A A C G C A A G T G G T T C T G G T A G C G T G A ...
 2530 2540 ...
 ...THR ALA LYS THR SER SER ASN VAL ASN ILE THR
 ... C T G C G A A A A C C T C A A G T A A T G T G A A T A T C A C C
 ...2550 2560 2570 2580

GLY ASP LEU SER THR ILE ASN GLY LEU ...
 G G G G A T T A A G C A C A A T A A T G G G T T A A ...
 2590 2600 ...
 ...ASN ILE ILE SER GLU ASN GLY ARG ASN THR VAL
 ... A T A T C A T T T C G G A A A A T G G T A G A A A C A C T G T G
 ...2610 2620 2630 2640

ARG LEU ARG GLY LYS GLU ILE ASP VAL ...
 C G C T T A A G A G G C A A G G A A A T T G A T G T G A ...
 2650 2660 ...
 ...LYS TYR ILE GLN PRO GLY VAL ALA SER VAL GLU
 ... A A T A T C C A A C C A G G T G T A G C A A G C G T A G A A
 ...2670 2680 2690 2700

135/235

FIG.24N

GLU VAL ILE GLU ALA LYS ARG VAL LEU ...
 GAGGTAAATTGAGCGAACAACGCGTCTTG...
 2710
 ...
 ...GLU LYS VAL LYS ASP LEU SER ASP GLU GLU ARG
 ...AAGAAAGTAAAGATTATTCTGACGAGAGAGA
 ...2730 2740 2750 2760

GLU THR LEU ALA LYS LEU GLY VAL SER ...
 GAACACTAGCCCAACCTTGGTGTAGTG...
 2770
 ...
 ...ALA VAL ARG PHE VAL GLU PRO ASN ALA ILE
 ...CTGTACGTTTCTGTTGAGCCCAATAATGCCATT
 ...2790 2800 2810 2820

THR ILE ASN THR GLN ASN GLU PHE THR ...
 ACGATTATAACACAATAAGATTACAA...
 2830
 ...
 ...THR ARG PRO SER SER GLN VAL ILE SER GLU
 ...CCAGACCGTCAAGTCAAGTGATTAATTCTGAA
 ...2850 2860 2870 2880

GLY LYS ALA CYS PHE SER SER GLY ASN ...
 GGTAAGCGGTGTTCTCAAGTGGTAATG...
 2890 2900
 ...

136/235

FIG.240

...GLY ALA ALA VAL CYS THR ASN VAL ALA ASP ASP
... GCGCAGCAGTATGTATGCTGACGAT 2940
...2910 2920 2930

GLY GLN PRO ***
GGACAGCCGTAG
2950

FIG.25A

PMH1 *hmm2A* sequence

LYS GLU TRP LEU LEU ASP PRO ASP ASP ...
 A A G A G T G G T T G T T A G A C C C G G A T G A T G...
 10 20
 ... VAL THR ILE ALA ALA GLY ALA PRO GLY ARG ASN
 ... T A A C T A T T G C C C G C A G G C C A G G A C G T A A C
 ... 30 40 50 60
 137/235
 ASP GLY SER VAL ASP ASP PHE PHE PRO ...
 G A T G G T T C A G T A G A C G A C T T T T T C C C A...
 70 80
 ... THR GLY ARG GLY ASP ASP ALA SER ASN ALA LYS
 ... C T G G A A G A G G G G A T G A T G C T A G T A A T G C A A A A
 ... 90 100 110 120
 THR ASN HIS PRO ASP LYS PRO THR LEU ...
 A C A A C C A T C C A G A C A A G C C G A C A T T A A...
 130 140
 ... THR ASN THR VAL GLU ASN ALA LEU LYS ASN
 ... C A A C A C A A C T G T T G A G A A C G C A T T A A A A A C
 ... 150 160 170 180

138/235

FIG.25B

```

ASN  THR  PHE  VAL  ASN  ILE  THR  ALA  LYS  ...
A  A  C  A  C  C  T  T  T  G  T  T  A  A  C  A  T  A  A  C  C  G  C  C  A  A  A  ...
190
...
...ASN  LYS  ILE  THR  VAL  ASN  SER  ASP  ILE  ASN  ILE
...  A  T  A  A  A  A  T  C  A  C  A  G  T  T  A  A  T  A  G  C  G  A  C  A  T  C  A  A  T  A  T  C
210
... 220
230
240

```

```

LYS  GLY  GLY  ALA  HIS  LEU  THR  LEU  TYR  ...
A  A  G  G  T  G  G  C  C  C  C  A  C  C  T  A  A  C  C  T  C  T  A  T  A  ...
250
...
...SER  LYS  ASN  LYS  LYS  LYS  SER  SER  VAL  LYS  ILE
...  G  C  A  A  A  A  C  A  A  T  A  A  A  A  A  A  G  T  A  G  C  G  T  T  A  A  G  A  T  T
270
... 280
290
300

```

```

ASN  GLY  ASN  ILE  THR  SER  THR  THR  ASN  ...
A  A  T  G  G  C  A  A  T  A  T  T  A  C  T  T  C  T  C  T  A  C  C  A  C  T  A  A  C  G  ...
310
...
...GLY  ASN  LEU  THR  ILE  TYR  SER  SER  GLY  TRP  VAL
...  G  A  A  A  C  T  T  A  A  C  T  A  T  T  T  A  C  T  C  C  A  G  C  G  G  C  T  G  G  G  T  T
330
... 340
350
360

```

```

ASP  ILE  HIS  LYS  ASN  ILE  THR  LEU  ASN  ...
G  A  T  A  T  C  C  A  T  A  A  A  A  C  A  T  T  A  C  G  C  T  T  A  A  C  A  ...
370
...
380
...

```

FIG.25C

139/235

...THR GLY TYR LEU ASN ILE THR ALA GLY GLY SER
 ... C A G G T T A C C T G A A T A T T A C C G C T G G G G T T C T
 ... 390 400 410 420

VAL ALA PHE GLU LYS ALA GLY ASN GLU ...
 G T A G C C T T C G A G A A A G C C G G A A A T G A G A...

430

440

...LYS GLY ARG GLN VAL SER GLU SER VAL ILE LYS
 ... A A G G G C C C A A G T A T C A G A A T C T G T A A T C A A A
 ... 450 460 470 480

ALA GLN GLY VAL ILE THR SER GLY VAL ...
 G C C C A G G G A G T T A T C A C C T C A G G T G T A G...

490

500

...GLY GLU GLY PHE ARG PHE ASN ASN VAL SER LEU
 ... G G G A A G G C T T T A G G T T T A A T A A C G T C T C C C T A
 ... 510 520 530 540

ASN GLY VAL GLY ALA GLY LEU ARG PHE ...
 A A T G G C G T T G G C G C A G G A C T G C G C T T C G...

550

560

...VAL GLY GLN LYS ASN ILE SER SER ASN SER TRP
 ... T T G G T C A G A A A A T A T C A G T A G C A A C T C T T G G
 ... 570 580 590 600

140/235

FIG.25D

ARG GLU ASN THR ILE LYS ASN ARG PHE ...
 A G A G A A A C A C C A T C A A A A C A G A T T C G ...
 610
 ...
 ...ASP GLY ASN LEU ASN ILE SER GLY LYS VAL ASN
 ... A T G G G A A T T A A A T A T C T C A G G A A A G G T A A A T
 620
 ... 630 640 650 660

VAL SER MET ASP VAL SER GLY THR LYS ...
 G T T C A A T G G A T G T A T C C G G G A C A A A G T ...
 670
 ...
 ...TRP HIS THR ARG ILE ASN GLY ARG THR TYR TRP
 ... G G C A T A C A G A A T T A C G G G C G C A C C T A C T G G
 680
 ... 690 700 710 720

ASN VAL THR THR LEU ASN VAL ALA SER ...
 A A T G T A A C C A C T C T A A A C G T T G C C T C A G ...
 730
 ...
 ...GLY SER SER PHE ASN LEU SER ILE ASP ALA SER
 ... G T A G T A G T T T C A A T C T C A G T A T C G A C G C C A G T
 740
 ... 750 760 770 780

GLY ILE SER SER GLY ASN GLN ASP ASP ...
 G G A T T T C T C A G G T A A C C A G G A C G A C A ...
 790
 ... 800

141/235

FIG.25E

...ILE THR ASN ARG GLY LEU ASN GLY ILE THR PHE
 ... T A C A A A T A G G G T T T A A T G G C A T A C A T T T
 ... 810 820 830 840

ASN GLY GLU ASN THR PHE ASN ILE ALA ...
 A A T G G A G A A A C A C T T T A A T A T C G C A C ...
 ... 850 860 ...
 ...GLN GLY SER THR ALA ASN PHE HIS ILE LYS THR
 ... A G G G C T C A A C A G C T A A C T T T C A T A T C A A A C G
 ... 870 880 890 900

SER VAL MET THR PRO LYS PRO ASN SER ...
 T C A G T A A T G A C C C C T A A C C C A A C T C G A ...
 ... 910 920 ...
 ...ASN TYR ALA LEU PHE ASN GLY ASN ILE SER VAL
 ... A C T A C G C A T T A T T T A A T G G A A A T A T T C A G T T
 ... 930 940 950 960

LEU GLY GLY GLY THR VAL ASN PHE GLU ...
 T T A G G A G G A G A C T G T C A A C T T T G A A C ...
 ... 970 980 ...
 ...LEU ASN ALA SER SER THR HIS THR THR SER
 ... T T A A T G C C T C A T C T A G C A C C C A C A C T T C T
 ... 990 1000 1010 1020

142/235

FIG.25F

```

GLY  ALA  ILE  ILE  ASN  SER  GLN  ASN  PHE  ...
GGCGCAATTATAATTCTCTCAAAATTTTA...
1030
...ASN  VAL  SER  GLY  GLY  SER  LYS  LEU  ASN  LEU  LYS
...ATGTCCTCAGGTGGGTCAAAATTAATCTCAAG
...1050 1060 1070 1080

ALA  SER  GLY  SER  THR  ASN  THR  ALA  PHE  ...
GCTTCAGGCTCAACAATAACCGCTTTT...
1090
...LEU  ILE  LYS  ASN  ASN  LEU  THR  LEU  ASN  ALA  THR
...TAATAAATAATAATTAACTTTAAACGCTACT
...1110 1120 1130 1140

GLY  GLY  ASN  ILE  GLU  ILE  LYS  GLN  VAL  ...
GGAGGTAATAATAGAAATAATAACAGGTTG...
1150
...GLU  GLY  THR  ASP  SER  ARG  ILE  GLN  LYS  GLY  VAL
...AGGGTACCGATTTCGGCGCATTCAAAAGGTGTT
...1170 1180 1190 1200

VAL  ALA  GLU  GLN  ASN  ILE  ILE  PHE  GLU  ...
GTAGCCGAACAACAATAATTTTGAAAG...
1210 1220

```

143/235

FIG.25G

...GLY GLY ASN ILE THR LEU GLY SER GLN LYS ALA
 ...GGGGTAACATCACCCCTTGGCTCCCAAAAGCC
 ...1230 1240 1250 1260

PRO THR GLU ILE LYS GLY ASP VAL THR ...
 CCAACAGAAATAAAGGCGATGTACCG...

1270

1280

...VAL LYS GLN GLY THR ASN ALA THR LEU ARG SER
 ...TCAACAAGGAACCAACGCCACTCTCAGAGC
 ...1290 1300 1310 1320

ALA ASN PHE ASP ASN HIS LYS GLY ALA ...
 GCGAATTGTGACACCAACAAGGTGCT...

1330

1340

...LEU ILE VAL ASN GLY ASN VAL THR ALA ASN GLY
 ...TAATTGTGAATGGAAACGTTACCGCCAATGGC
 ...1350 1360 1370 1380

ASN LEU THR ALA ASP GLY ASP THR ILE ...
 AACCTTAC TGCGGACGGCGACACTATA...

1390

1400

...LYS ILE LYS GLY ASN LEU ASP VAL ALA GLN GLY
 ...AAATAAAGGCAATCTTGATGTGCAACAAGGC
 ...1410 1420 1430 1440

144/235

FIG.25H

ALA LYS PHE ASN GLY SER THR LYS ASN ...
 G C T A A A T T T A C G G C A G C A C A A A A C A ...
 1450
 ...
 ...ASN LEU ASN ILE THR GLY THR PHE THR ASN ASN
 ... A C C T A A A C A T T A C T G G C A C C T T T A C C A C A C
 1460
 ...1470
 1480
 1490
 1500

GLY THR SER ILE ILE ASP ILE THR GLN ...
 G G C A C T T C T A T A T C G A T A T A C A C A G ...
 1510
 ...
 ...GLY VAL VAL ASN LEU GLY ASN VAL THR ASN ASP
 ... G G T G G T A A A C C T T G G T A A T G T T A C C A A T G A C
 1520
 ...1530
 1540
 1550
 1560

GLY LYS LEU ASN ILE THR THR HIS ALA ...
 G G C A A A T T A A A C A T C A C C A C T C A C G C C A ...
 1570
 ...
 ...LYS SER GLY GLN LYS SER ILE ILE ARG GLY ASP
 ... A G A G C G G T C A A A A A G C A T T A T C C G C G G A G A T
 1580
 ...1590
 1600
 1610
 1620

ILE ILE ASN LYS GLN GLY ASN LEU ASN ...
 A T A A T T A A C A A C A G G G A A T T A A A T A ...
 1630
 1640
 ...

145/235

FIG.25I

...ILE THR ASP ASN ASN SER ASN ALA GLU ILE GLU
 ... T T A C G G A C A A T A A T A G T A A T G C T G A A A T T G A A
 ...1650 1660 1670 1680

ILE GLY GLY ASN ILE SER GIN LYS GLU ...
 A T T G G C G G C A A T A T C T C G C A A A A G A G ...
 1690 1700
 ...GLY ASN LEU THR ILE SER SER ASP LYS VAL ASN
 ... G T A A T C T C A C C A T T T C T T C T G A T A A G T C A A T
 ...1710 1720 1730 1740

ILE THR LYS GIN ILE THR ILE LYS ALA ...
 A T T A C C A A A C A G A T A C A A T C A A A G C A G ...
 1750 1760
 ...GLY VAL ASP GLY GLU SER SER SER SER THR
 ... G C G T T G A T G G G G A G A G T T C T A G T T C A G C A C A
 ...1770 1780 1790 1800

ALA SER ASP ALA ASN LEU THR ILE LYS ...
 G C A A G T G A T G C C A A T C T A A C C A T T A A A ...
 1810 1820
 ...THR LYS GLU LEU THR PHE THR ASP ASN LEU ASN
 ... C C A A A G A G T T A A C A T T C A C A G A C A A T C T A A A C
 ...1830 1840 1850 1860

146/235

FIG.25J

```

ILE SER GLY PHE ASN LYS ALA GLU ILE ...
A T T C A G G T T T A A T A A G C A G A A T T A ...
1870
...
...THR ALA LYS ASP ASN SER ASP LEU ILE GLY
... C A G C T A A A G A T A C A G T G A T T A A T T A T T G G C
...1890 1910 1920

```

```

LYS ALA SER SER ASP ASN SER ASN ALA ...
A A G G C T A G C A G T G A C A A C A G T A A T G C T A ...
1930
...
...LYS GLN VAL THR PHE ASP LYS VAL LYS ASP SER
... A A C A A G T A A C C T T T G A C A A G G T T A A A G A T T C A
...1950 1970 1980

```

```

LYS ILE SER ALA GLY ASN HIS ASN VAL ...
A A A T C T C A G C T G G C A A T C A C A A T G T A A ...
1990
...
...THR LEU ASN SER LYS VAL GLU THR SER ASN SER
... C A C T A A A T A G C A A A G T G G A A A C G T C T A A T A G C
...2010 2030 2040

```

```

ASP GLY SER SER THR GLY ASN GLY SER ASP ...
G A T G G T A G C A C C G G A A A C G G T A G C G A T G ...
2050 2060

```

FIG.25K

147/235

...ASP ASN ASN ILE GLY LEU THR ILE SER ALA LYS
 ... A C A C A A T A T C G G C T T A A C T A T T C C G C A A A
 ...2070 2080 2090 2100

ASP VAL THR VAL ASN SER ASN ILE THR ...
 G A T G T A A C G G T A A A T A G T A A T C A C C T ...
 2110 2120
 ...SER HIS LYS THR VAL ASN ILE SER ALA SER GLU
 ... C T C A C A A A C A G T A A A T A T C T C T G C A T C A G A A
 ...2130 2140 2150 2160

GLY GLY ILE THR THR LYS ALA GLY THR ...
 G G A G G T A T C A C T A C T A A A G C A G G C A C A A ...
 2170 2180
 ...THR ILE ASN ALA THR THR GLY SER VAL GLU VAL
 ... C C A T T A A T G C G A C C A C A G G T A G C G T G G A A G T A
 ...2190 2200 2210 2220

THR ALA LYS THR GLY ASP ILE SER GLY ...
 A C T G C T A A A A C A G G C G A T A T T A G C G G T A ...
 2230 2240
 ...THR ILE SER GLY LYS THR THR VAL SER VAL THR ALA
 ... C G A T T T C C G G T A A G A C A G T A A G T G T T A C A G C A
 ...2250 2260 2270 2280

09/806709

148/235

FIG.25L

```

SER  THR  GLY  ASP  LEU  THR  VAL  ARG  LYS  ...
AGCACTGGCGATTTAACCTGTTAGGAAG...
2290
...
...ALA  ALA  THR  ILE  SER  VAL  THR  GLU  GLY  ALA  ALA
...CTGCAACCAATTAGTGTCAGACAGAGGAGCTGCA
...2310                2330                2340

THR  LEU  THR  ALA  THR  GLY  ASN  THR  LEU  ...
ACCTTAACCGCAACACAGGGAATACCTTGA...
2350
...
...THR  THR  GLU  ALA  GLY  SER  SER  ILE  THR  SER  THR
...CTACTGAAGCCGGTCTAGCATCTCACT
...2370                2390                2400

LYS  GLY  GLN  VAL  ASP  LEU  SER  ALA  GLN  ...
AAGGTCAGGTAGACCTTTTCAGCTCAGG...
2410
...
...ASP  GLY  SER  ILE  ALA  GLY  GLN  ILE  SER  ALA  ALA
...ATGGTAGCATTGCAAGGACAAATTAGTGCAGCT
...2430                2440                2450                2460

ASN  VAL  THR  LEU  ASN  THR  THR  GLY  THR  ...
AATGTGACATTATAATACCAAGGCCACT...
2470                2480
...
```

149/235

FIG.25M

...LEU THR THR VAL GLU GLY SER ASN ILE LYS ALA
 ... T A A C T A C T G T A G A A G G T T C A A C A T T A A G G C A
 ...2490 2500 2510 2520

THR SER GLY THR LEU ALA ILE ASN ALA ...
 A C C A G T G G C A C C T T A G C T A T T A A C G C A A ...
 2530 ...
 ...LYS ASP ALA LYS LEU ASP GLY THR ALA SER GLY
 ... A A G A C G C T A A G C T A G A T G G T A C G G C A T C A G G T
 ...2550 2560 2570 2580

ASN ARG THR GLU VAL ASN ALA THR ASN ...
 A A C C G T A C A G A A G T A A T G C A A C T A A C G ...
 2590 ...
 ...ALA SER GLY SER GLY SER VAL THR ALA LYS THR
 ... C A A G T G G T T C T G G T A G C G T G A C T G C G A A A C C
 ...2610 2620 2630 2640

SER SER ASN VAL ASN ILE THR GLY ASP ...
 T C A A G T A A T G T G A A T A T C A C C G G G A T T ...
 2650 ...
 ...LEU SER THR ILE ASN GLY LEU ASN ILE ILE SER
 ... T A A G C A C A A T A A A T G G G T T A A A T A T C A T T T C G
 ...2670 2680 2690 2700

150/235

FIG.25N

```

GLU ASN GLY ARG ASN THR VAL ARG LEU ...
G A A A T G G T A G A A A C A C T G T G C G C T T A A ...
2710
...
...ARG GLY LYS GLU ILE ASP VAL LYS TYR ILE GLN
... G A G G C A A G G A A A T T G A T G T G A A A T A T A T C C A A
...2730 2740 2750 2760

PRO GLY VAL ALA SER VAL GLU GLU VAL ...
C C A G G T G T A G C C A G C G T A G A A G A G G T A A ...
2770
...
...ILE GLU ALA LYS ARG VAL LEU GLU LYS VAL LYS
... T T G A A G C G A A A C G C G T C C T T G A G A A A G T A A A A
...2790 2800 2810 2820

ASP LEU SER ASP GLU GLU ARG GLU THR ...
G A T T A T C T G A C G A A G A A A G A G A A A C A C ...
2830
...
...LEU ALA LYS LEU GLY VAL SER ALA VAL ARG PHE
... T A G C C A A A C T T G G T G T A G T G C T G T A C G T T C
...2850 2860 2870 2880

VAL GLU PRO ASN ASN ALA ILE THR ILE ...
G T T G A G C C A A A T A A T G C C A T T A C G A T T A ...
2890
...
```

FIG.250

151/235

...ASN THR GLN ASN GLU PHE THR THR ARG PRO SER
 ... ATACACAATAATTTACACAGACCGTCA
 ...2910 2920 2930 2940

SER GLN VAL ILE ILE SER GLU GLY LYS ...
 AGTCAGTGATAATTCTCTGAGGTAAAG...

2960

2950

...ALA CYS PHE SER SER GLY ASN GLY ALA VAL
 ... CGTGTTTCTCAAGTGGTAATGGCGCAGCAGTA
 ...2970 2980 2990 3000

CYS THR ASN VAL ALA ASP GLY GLN ...
 TGTAACCAATGTTGCTGACGATGGACAGC...
 3010 3020 ...
 ...PRO ***
 ...CGTAG
 ...3030

FIG.26A

Strain 15 *hmm1A* sequence

152/235

```

  LYS  GLU  TRP  LEU  LEU  ASP  PRO  ASP  ASN  VAL  ...
  A A G A G T G T T A G A C C C G G A T A A T G T A A ...
  10
  ...
  20
  ...
  30
  ...
  ...THR  ILE  GLU  ALA  PRO  SER  TYR  SER  ARG  GLY
  ... C A A T T G A A G C C C C T T C C C T A T T C T C G C G G T
  ...
  40
  ...
  ASN  ALA  GLY  ILE  ASP  SER  GLU  PHE  PRO  GLY  ...
  A A T G C C G G T A T A G A T A G T G A A T T C C C G G G C G ...
  70
  ...
  80
  ...
  ...GLY  SER  GLY  THR  LYS  GLU  SER  PRO  LYS  THR
  ... G T T C G G G C A C A A A G G A A A G C C C T A A A C A
  ...
  100
  ...
  ASN  GLY  GLU  GLN  PRO  THR  VAL  LEU  THR  ASN  ...
  A A C G G C G A A C A G C C G A C A G T A T T A C C A A T G ...
  130
  ...
  140
  ...
  ...GLU  THR  ILE  SER  ASN  TYR  LEU  LYS  SER  GLY
  ... A A A C C A T T T C A A A T T A T C T G A A A G C G G C
  ...
  160
  ...
  170
  ...
  180

```

153/235

FIG.26B

```

THR TRP VAL MET ASN ILE THR ALA LYS LYS ...
ACCTGGGTAAATGAATAATACAGCCAGAA...
190      200      210 ...
...ASN LEU THR VAL ASN SER ILE ASN ILE
...ATCTTACCGTTAACAGCTCAATTACAATT
...      220      230      240

GLY ASP SER SER HIS LEU ILE LEU HIS SER ...
GGAGACAGCTCCCACTTAATCCTTCA TAGTG...
250      260      270 ...
...GLU GLY LYS ASN ASN GLY VAL LYS ILE
...AAGGCAAGAAATAACGGCGGTGAAGATT
...      280      290      300

LYS GLU ASP ILE THR SER ASN GLY GLY ASN ...
AAGAGACATTAACCTCTAATGGCGGAAC...
310      320      330 ...
...LEU THR ILE GLN SER GLY TRP VAL ASP
...TAACCATTCATAACGGCGGATGGTTGAT
...      340      350      360

VAL HIS LYS ASN ILE THR LEU GLY THR GLY ...
GTTCACAAAATAATTACGCTTGGCACAGGCA...
370      380      390 ...

```

154/235

FIG.26C

...THR LEU ASN ILE THR ALA LYS GLY SER ILE
 ... CCTTGAAATATTACAGCTAAAGGATCCATA
 ... 400 410 420

ALA PHE GLU GLY ASN GLY THR GLU LYS ALA ...
 GCCTTTGAGGGAACGGGTACAGAAAGAGCC...
 430 440 450 ...
 ...ARG ASN ALA SER SER ALA GLN ILE THR ALA
 ... GCACGCATCAAGCGCTCAATAACCGCG
 ... 460 470 480

GLN GLY THR ILE THR ASN THR GLY ASP GLN ...
 CAGGGAACATAACCAATACTGGCGATCAAA...
 490 500 510 ...
 ...LYS GLN LEU ARG LEU ASN ASN VAL SER ILE
 ... AACAACTCAGACTTAATAATGTATCTATT
 ... 520 530 540

ASN GLY THR GLY ILE GLY LEU ASN PHE VAL ...
 AATGGGACGGGTATAGGTTTAAATTTGTTT...
 550 560 570 ...
 ...SER ILE GLN PRO ASN THR SER HIS ARG PHE
 ... CAATTCAGCCCTAACACTTCTCAGATT
 ... 580 590 600

155/235

FIG.26D

```

ASP  GLY  GLU  LEU  ILE  ILE  SER  GLY  ARG  VAL  ...
GATGGGAGCTTATTATTTCAGGGAGAGTAC...
610
...HIS  VAL  ASN  GLN  THR  THR  PRO  LYS  ASN  LEU
...ATGTTAATCAAAACCACACCTAAACCTG
640
...
630
650
660

```

```

SER  PHE  TRP  LYS  VAL  SER  ASP  GLU  SER  TYR  ...
TCTTTTGGGAAGGTATCCGATGATCTTATT...
670
...TRP  ASN  VAL  SER  HIS  LEU  THR  VAL  LYS  GLU
...GGAATGTCAGCCATCTTACCGTAAAGAG
700
...
710
720

```

```

LYS  SER  ALA  PHE  SER  PHE  THR  LYS  PHE  ALA  ...
AAGTCAGCATTCATCTCATTTACCAAGTTTGGGT...
730
740
...LEU  ASN  ASN  ASN  HIS  GLY  ARG  GLU  THR  SER
...TAAATAACAATCATGGCCGAGAGACTTCC
760
...
770
780

```

```

ARG  TYR  ARG  LYS  GLY  GLY  VAL  ILE  PHE  ...
AGATACCGCAAGGTGGAGGTGTATCTTTC...
790
800
810
...

```

156/235

FIG.26E

...ARG SER PRO THR GLY HIS THR ASN PHE THR
 ... GCTCACCCTACCGGTCAACAATAATTCACA
 ... 820 830 840

VAL LYS GLN GLY SER VAL ALA ASN PHE SER ...
 GTTAAACAAGGCTCAGTGCGCTAATTTTCTAT...
 ... 850 860 870 ...
 ...PHE LYS ALA LYS ASN ASP THR ASN HIS ALA
 ... TCAAGGC AAAAATGATACAAATCATGCA
 ... 880 890 900

ASN GLN LEU PRO ILE GLN PHE ASN SER ASN ...
 AATCAACTCCCGATTTCAGTTTAACTCTATAA...
 ... 910 920 930 ...
 ...ILE SER VAL ASP GLY GLY LYS VAL LEU
 ... TCTCAGTCCGATGGAGGAGGGAAGTCCCTT
 ... 940 950 960

PHE CYS ILE THR SER ASN TYR SER GLY ARG ...
 TTTTGTAATAACCTCCCAACTACTCCGGCAGAT...
 ... 970 980 990 ...
 ...SER VAL GLY ILE GLY MET SER SER ILE ASN
 ... CAGTGGGATAGGAATGTC TAGCATTAAT
 ... 1000 1010 1020

157/235

FIG.26F

VAL SER ASP GLY SER ASN LEU THR PHE ASN ...
 GTTCTGATGGCTCAACCTTACTTTTAAAT...
 1030
 ...SER SER ILE ARG GLY GIN GLU ALA PHE ASN
 ...CTTCCATTTCGGCGCCAGGAAGCCTTTAAAT
 1040
 ... 1060 1070 1080

ILE SER LYS ASP LEU THR ILE ASN ALA THR ...
 ATCAGTAAGAATTTAACCAATAATGCCACCG...
 1090
 ...GLY SER PHE PHE GLU LEU GLY GIN TYR SER
 ...GTTCAATTTTGTGAACCTTGGGCAATACTCG
 1100 1110 1120 1130 1140

ASP THR PHE ASN GLY ASN GLY PHE ASN HIS ...
 GATACCTTTAATGGTAAATGGCTTTAACCAACG...
 1150
 ...ASP ALA ILE LYS SER THR HIS ASN ILE SER
 ...ACGCCATTAAATCACTCAACAATAATATCC
 1160 1170 1180 1190 1200

ILE LEU GLY GLY ASN VAL THR LEU GLY GLY ...
 ATCTTAGGTGGCAATGTTACCCCTTGGCGGGC...
 1210 1220 1230 ...

158/235

FIG.26G

...GLN ASP SER SER THR ILE THR GLY ASN
 ... A A G A T T C A A G C A G T A C C A T T A C A G G T A A T
 ... 1240 1250 1260

ILE ASN ILE SER GLN ALA ALA ASN VAL THR ...
 A T C A A T A T C T C T C A G G C A G C A A A T G T T A C C T ...
 1270 1280 1290 ...
 ...LEU ARG ALA TYR ASN GLY ASN GLY ARG ASN
 ... T G C G A G C T T A T A T G G T A A C G G T C G A A A C
 ... 1300 1310 1320

LYS GLN LEU THR LEU GLY ASN VAL SER ILE ...
 A A C A A C T A A C C C T T G G C A A T G T A T C T A T T G ...
 1330 1340 1350 ...
 ...GLU GLY ASN LEU SER LEU ILE GLY ALA SER
 ... A A G G G A A T T A A G T T T A A T C G G T G C A A G T
 ... 1360 1370 1380

ALA ASN ILE ASN GLY ASN LEU SER VAL LYS ...
 G C A A A T A T T A A C G G C A A C C T T T C C G T T A A A G ...
 1390 1400 1410 ...
 ...GLU ASN ALA LYS PHE LYS GLY GLU THR GLN
 ... A A A A T G C T A A A T T T A A A G G G G A A C C C A A
 ... 1420 1430 1440

159/235

FIG.26H

ASP ASN LEU ASN ILE THR GLY THR PHE ILE ...
 G A C A A C T T G A A C A T C A C C G G C A C C T T T A T C A ...
 1450
 ...ASN ASN GLY ASP SER LYS ILE ASN ILE SER
 ... A T A A C G G C G A C T C T A A A A T C A A T A T A T C T
 1460 1470 ...
 ... 1480 1490 1500

GLN GLY VAL VAL LYS LEU GLY ASN VAL THR ...
 C A G G A G T G G T A A A A C T T G G C A A T G T T A C C A ...
 1510 1520 1530 ...
 ...ASN ASP GLY ASP LEU ASN ILE THR THR HIS
 ... A T G A T G G T G A T T T A A C A C A T T A C C A C T C A C
 1540 1550 1560

ALA LYS HIS ASN GLN ARG SER ILE ILE GLY ...
 G C T A A C A C A C C A A A G A A G C A T C A T C G G C G ...
 1570 1580 1590 ...
 ...GLY ASP ILE ILE ASN LYS LYS GLY SER LEU
 ... G A G A T A T A T C A C A C A A A A A G G A A G C T T A
 1600 1610 1620

ASN ILE THR ASP SER ASN LYS ASN ALA GLU ...
 A A T A T C A G A C A G T A A T A A G A A T G C T G A A A ...
 1630 1640 1650 ...

FIG.26I

160/235

...ILE GLN ILE GLY GLY ASN ILE SER GLN LYS
 ... TCCAAAT TGGCGGCAATATCTCGCAAA A A
 ... 1660 1670 1680

GLU GLY ASN LEU THR ILE SER SER ASP LYS ...
 GAGGCAATCTCACGATTCTCTCCGATAA A A...
 1690 1700 1710 ...
 ...ILE ASN ILE THR ASN GLN ILE THR ILE LYS
 ... TCAATATTACCAATCAGATAACAATCA A A
 ... 1720 1730 1740

ALA GLY VAL ASP GLY GLU ASN SER ASP SER ...
 GCAGGTGTTGATGGGGAGAAATTCGGATTCA G...
 1750 1760 1770 ...
 ...ASP ALA THR ASN ASN ALA ASN LEU THR ILE
 ... ACGCGACAAACAATGCCAATCTAACCA TT
 ... 1780 1790 1800

LYS THR LYS LYS GLU LEU LYS LEU THR GIN ASP ...
 AAAACCAAGAAATTGAAATTAACGCAAGACC...
 1810 1820 1830 ...
 ...LEU ASN ILE SER GLY PHE ASN LYS ALA GLU
 ... TAAATATTTCAGGTTTCAATAAAGCAGAG
 ... 1840 1850 1860

161/235

FIG.26J

```

ILE THR ALA LYS ASP GLY SER ASP LEU THR ...
A T T A C A G C T A A A G A T G G T A G T G A T T T A A C T A ...
1870
... ILE GLY ASN THR ASN SER ALA ASP SER THR
... T T G G T A A C A C C A A T A G T G C T G A T A G T A C T
1880
... 1900
... 1910
... 1920

```

```

ASN VAL LYS LYS VAL THR PHE ASN GLN VAL ...
A A T G C C A A A A A G T A A C C T T T A A C C A G G T T A ...
1930
... LYS ASP SER LYS ILE SER ALA GLY ASP HIS
... A A G A T T C A A A A A T C T C T G C T G G C G A C C A T
1940
... 1950
... 1960
... 1970
... 1980

```

```

ASN VAL THR LEU ASN SER LYS VAL GLU THR ...
A A T G T G A C A C T A A A T A G C A A A G T G G A A A C A T ...
1990
... SER GLY ASN THR ASP ASN THR GLY ASP GLY
... C T G G T A A T A C T G A C A C A C T G G A G A C G G C
2000
... 2010
... 2020
... 2030
... 2040

```

```

SER GLY ASN ASN ALA GLY LEU THR ILE ALA ...
A G T G G C A A T A A T G C C G C T T A A C T A T T G C C G ...
2050
... 2060
... 2070
...

```

162/235

FIG.26K

...ALA LYS ASN VAL GLU VAL LYS ASN ASN ILE
 ... CGA A A A T G T A G A A G T A A A A A C A C A T T 2100
 ... 2080 2090

THR SER ASN LYS THR THR VAL ASN ILE THR ALA ...
 A C T T C T A A C A A A C A G T A A A T A T C A C C G C G T ...
 2110 2120 ... 2130 ...
 ...SER GLU LYS LEU THR THR LYS ALA ASP ALA
 ... C A G A A A A C T T A C C A C C A A A G C G G A T G C A 2160
 ... 2140 2150

THR ILE ASN ALA THR THR GLY ASN VAL GLU ...
 A C C A T T A A T G C A A C C A C T G G T A A C G T A G A A G ...
 2170 2180 ... 2190 ...
 ...VAL THR ALA LYS THR GLY ASP ILE LYS GLY
 ... T G A C A G C C A A A C A G G T G A T A T T A A A G G T 2220
 ... 2200 2210

GLU VAL LYS SER THR THR SER GLY ASN VAL ASN ...
 G A G T C A A A T C C A C T T C C G G T A A T G T A A A T A ...
 2230 2240 ... 2250 ...
 ...ILE THR ALA ASN GLY ASP THR LEU ASN VAL
 ... T T A C A G C A A A C G G C G A C A C G C T T A A T G T A 2280
 ... 2260 2270

163/235

FIG.26L

```

SER ASN VAL SER GLY ASN ALA VAL THR ILE ...
AGT AATGTTTCAGGCAATGCTGTATCCATCA...
2290      2300      2310 ...
...THR ALA ASP LYS GLY LYS LEU THR THR GLN
... CTGCAGATAAGGGCAATAATTAACCAACCACA
...      2320      2330      2340

```

```

ALA SER SER SER ILE THR SER ASN ASN GLY ...
GCAAGCTCTAGCATTACCTCTCAACAATGGCC...
2350      2360      2370 ...
...GLN THR THR LEU THR ALA LYS ASP GLY SER
... AGACAACCTCTTACAGCCCAAGGATGGCAGT
...      2380      2390      2400

```

```

ILE ALA GLY SER ILE ASN ALA ALA ASN VAL ...
ATCGCAGGAAGCATCATGCGCCCAATGTGA...
2410      2420      2430 ...
...THR LEU ASN THR THR GLY THR LEU THR THR
... CATTAAATACCAAGGCACCTTAACTACT
...      2440      2450      2460

```

```

VAL GLU GLY SER ASN ILE ASN ALA ALA SER ...
GTAGAAGGTTCAACAATTACGCAAGCAGTG...
2470      2480      2490 ...

```

164/235

FIG.26M

...GLY THR LEU VAL ILE ASN ALA LYS ASP ALA
 ...GTACCTTGGTTATTATGCAAAAGATGCT
 ... 2500 2510 2520

LYS LEU ASN GLY ALA SER GLY ASP HIS ...
 AAGTTGAACGGCGCGCATCAGGTGACCA...
 2530 2540 2550 ...
 ...THR VAL VAL ASN ALA THR ASN ALA SER GLY
 ...CAGTAGTAATAAGCAACTAACGCAAGTGGC
 ... 2560 2570 2580

SER GLY SER VAL THR ALA VAL THR SER SER ...
 TCTGGTAGTGTA CTGCGGTAACTCAAGTA...
 2590 2600 2610 ...
 ...ASN VAL ASN ILE THR GLY ASP LEU SER THR
 ...ATGTGAATAATCACCGGGGATTTAAGTACA
 ... 2620 2630 2640

VAL ASN GLY LEU ASN ILE SER LYS ASN ...
 GTAATAAGGATTAAATAATTCGAAATAAG...
 2650 2660 2670 ...
 ...GLY ARG ASN THR VAL VAL LEU LYS GLY THR
 ...GTAGAAACACCGTAGTGTTAAAGGTACT
 ... 2680 2690 2700

165/235

FIG.26N

```

GLU  ILE  GLU  VAL  LYS  TYR  ILE  GLN  PRO  GLY  ...
G A A T T G A G G T G A A T A T A T C C A G C C A G G T G ...
2710
...VAL  VAL  ALA  SER  VAL  GLU  GLU  VAL  ILE  GLU  ALA
...  T A G C A A G T G T A G A G A G A G T A A T T G A A G C G
...  2740  2750  2760

LYS  ARG  VAL  LEU  GLU  LYS  VAL  LYS  ASP  LEU  ...
A A C G C G T C C T T G A G A A A G T G A A A G A T T T A T ...
2770
...SER  ASP  GLU  GLU  ARG  GLU  THR  LEU  ALA  LYS
...  C T G A T G A A G A A A G A G A A A C A T T A G C T A A A
...  2800  2810  2820

LEU  GLY  VAL  SER  ALA  VAL  ARG  PHE  ILE  GLU  ...
C T T G G T G T A A G T G C T G T A C G T T T T A T T G A A C ...
2830
...PRO  ASN  ASN  THR  ILE  THR  VAL  ASN  THR  GLN
...  C A A T A A T A C C A T T A C G G T T A C A C A C A A
...  2860  2870  2880

```

166/235

FIG.260

```

ASN  GLU  PHE  THR  THR  ARG  PRO  SER  SER  GLN  ...
AATGAGTTTACACCAAGACCAATCAAGTCAAG...
2890
...VAL  THR  ILE  SER  GLU  GLY  LYS  ALA  CYS  PHE
...TGACAAATTCTGTGAAGGTAAAGCGCTGTTTC
...
2920
2930
2940

SER  SER  GLY  ASN  GLY  ALA  ALA  VAL  CYS  THR  ...
TCAAGTGGTAATGGCGCAGCAGTATGTACCA...
2950
...ASN  VAL  ALA  ASP  ASP  GLY  GLN  GLN  ***
...ATGTTGCTGACGATGGACAGCAGTAG
...
2980
2990

```

09/806709

FIG.27A

NTHi strain 15 *hmm2A* sequence

↓

```

ASN  SER  ALA  SER  GLY  SER  HIS  MET  PRO  ...
G A A T T C G G C T T C G G G A T C C C A T A T G C C G ...
10                                     20
...  GLU  ASN  VAL  TYR  ILE  ASN  ALA  GLY  ASP  ALA  GLY
...  G A G A A T G T A T A T A T A T G C A G G A G A C G C A G G
...  30                                     40                                     50                                     60
167/235

ARG  SER  ASP  THR  ASN  LEU  GLU  ASN  GLU  ...
G C G T A G T G A C A C T A A T T T A G A A A A C G A A ...
70                                     80
...  GLU  TYR  THR  GLY  THR  GLY  GLU  SER  ALA  ASP  THR
...  G A A T A C A C A G G A A C A G G A G A G T G C T G A T A C
...  90                                     100                                     110                                     120

PRO  LYS  ARG  ASN  ASN  ASN  THR  LYS  THR  ...
T C C A A A C G A A A C A A T A A C A C A A A G A C A ...
130                                     140
...  THR  LEU  THR  ASN  SER  THR  LEU  GLU  LYS  ILE  LEU
...  A C A C T A A C A A A C T C A A C G C T T G A G A G A T A T T
...  150                                     160                                     170                                     180

```

168/235

FIG.27B

```

ALA ARG GLY SER PHE VAL ASN ILE THR ...
AGCAAGAGGCTCTTTGTTTATAATCACT...
190                                     200
... ALA ASN ASN GLU ILE ARG VAL ASN SER ASP ILE
... GCCAACAAATGAATCAGAGTTAATAGTGATAT 240
... 210                                     220
ASN ILE GLY GLY ASN SER HIS LEU THR ...
CAATATCGGAGGCAACTCCCACTAACCC...
250                                     260
... LEU TRP SER SER LYS ASN LYS ASN SER GLY VAL
... CTC TGGAGCAGCAAAATAAACAAGTGCGCT 300
... 270                                     280
LEU ILE ASN GLY ASN ILE THR SER THR ...
TCTGATTATA TGGCAATA TCACTTCTACT...
310                                     320
... ALA ASN GLY ASN LEU THR ILE TYR SER SER GLY
... GCTAACGGAAACCTTAACCAATTACTCTAGCGG 360
... 330                                     340
TRP VAL ASP ILE HIS LYS ASN ILE THR ...
ATGGGTTGATTATTCATAAATAATTACG...
370                                     380
...

```

FIG.27C

... LEU GLU SER GLY ARG LEU ASN ILE THR THR LYS
 ... C T T G A A T C A G G A C G C T T A A A C A T T A C A C T A A
 ... 390 400 410 420

GLU GLY ASP VAL ALA PHE GLU LYS GLY ...
 A G A A G G A G A T G T C G C C T T T G A A A A G G G ...

430

440

... ASN ASN LEU THR ILE THR GLY GLN GLY THR THR ILE
 ... A A T A A C C T A A C C A T T A C A G G T C A A G G A C T A T
 ... 450 460 470 480

169/235

THR ALA GLY ASN ASN LYS GLY PHE ARG ...
 T A C A G C A G G C A A T A A T A A A G G C T T T A G A ...

490

500

... PHE GLU ASN VAL SER LEU ASN GLY THR GLY THR
 ... T T T G A A A A T G T C T C T C T A A A T G G C A C T G G G A C
 ... 510 520 530 540

GLY LEU LEU PHE ASN LEU SER ARG PRO ...
 T G G C T T G C T T T T A A T C T C A G T A G A C C A ...

550

560

... GLN LYS ASN ASN SER LEU VAL THR ASN TYR PHE
 ... C A A A A A C A A T A G T C T C G T C A C A A C T A T T
 ... 570 580 590 600

FIG.27D

```

ASN  GLY  THR  LEU  ASN  ILE  SER  GLY  SER  ...
TAA  TGG  GAC  TTT  AAA  TAT  TTT  CAG  GAG  GC...
610
...  VAL  ASN  ILE  SER  MET  ILE  PRO  PRO  ASN  ALA  THR
...  GTA  ATA  TCT  CAA  TGA  TTT  CCC  ACC  TAA  TGC  TAC
630
SER  ASN  TRP  TYR  SER  ARG  TYR  LYS  GLY  ...
AAG  CAA  TTG  GTAC  AGC  AGAT  ACA  AAG  GG...
670
...  ARG  THR  TYR  TRP  ASN  ILE  THR  HIS  LEU  ASN  ALA
...  CGA  ACC  TAT  TGG  AAT  ATA  ACC  CACT  TAA  ATGC
690
SER  GLU  ASP  SER  ASN  PHE  ASN  LEU  THR  ...
CTC  CGA  AGAT  AGCA  ACT  TTA  ACC  TTA  CT...
730
...  ILE  ASP  SER  SER  ALA  GLU  ASP  GLY  SER  ALA  PRO
...  ATT  GACT  CCT  CGG  CAG  AGG  ATGG  CTT  CAG  CCC
750
LEU  LEU  SER  SER  TYR  THR  LEU  ASN  GLY  ...
TCT  TTT  ATCC  AGTT  ATAT  ACC  TTAA  ACG  GC...
790
...
800

```

FIG.27E

171/235

... ILE SER PHE THR THR ASP THR PHE ASN VAL
 ... A T A T C A T T C A C C A G A T A C C A C C T T T A A T G T
 ... 810 820 830 840

ASN LYS ASN ALA LYS VAL ASN PHE ASN ...
 T A A T A A A A T G C A A A G T C A A C T T T A A C ...

850

860

... ILE LYS ALA PRO ILE GLY THR ILE ASN GLN TYR
 ... A T C A A A G C A C C A A T A G G G A C T A T A A T C A A T A
 ... 870 880 890 900

ASN ASN LEU ASN TYR ALA LEU PHE ASN ...
 C A A T A A C C T G A A T T A C G C A T T A T T C A A T ...

910

920

... GLY ASN ILE SER VAL SER GLY GLY GLY ASN VAL
 ... G G G A A C A T T T C A G T T T C A G G A G G G G A A T G T
 ... 930 940 950 960

THR PHE ARG LEU ASN ALA SER SER ...
 C A C C T T C A G G C T T A A C G C T T C A T C C T C T ...

970

980

... ASN GLN GLN THR PRO GLY VAL ILE ILE ASN SER
 ... A A C C A G C A A A C C C C T G G C G T A A T T A T A A T T C
 ... 990 1000 1010 1020

FIG.27F

LYS HIS LEU ASN ALA SER LYS GLY SER ...
 T A A C A C C T T A A T G C T T C A A A A G G G T C G ...
 1030 1040
 ... SER LEU ARG PHE ...
 ... A G C T T A A G A T T T G A A A C T A C A G G T T C A C A A A
 ...1050 1060 1070 1080

172/235

VAL GLY PHE LEU ILE ASN ASN ASP LEU ...
 A G T C G G T T T T A A T A A A T A T G A T T A ...
 1090 1100
 ... THR LEU ASN ALA THR GLY GLY ASN ILE SER LEU
 ... A C T T T A A A C G C C A C T G G A G G C A A T A T A T C G C T
 ...1110 1120 1130 1140

LEU GLN VAL GLU GLY ILE ASP GLY MET ...
 C T T G C A G G T T G A A G G C A T T G A C G G G A T G ...
 1150 1160
 ... ILE GLY GLU GLY VAL VAL ALA LYS LYS ASN ILE
 ... A T T G G T G A A G G C G T T G T A G C T A A A A A A C A T
 ...1170 1180 1190 1200

THR PHE THR GLY GLY ASN ILE THR PHE ...
 A A C C T T T A C T G G A G G C A A T A T C A C C T T T ...
 1210 1220

FIG.27G

... GLY SER LYS LYS ALA ILE THR GLU ILE LYS GLY
 ... GGCTCCAGAAAGCCATAACAGAAATCAAGG
 ...1230 1240 1250 1260

ASN VAL THR ILE ASN GLU ASN THR ASN ...
 CAATGTTACTATCAATGAACAACCAAC...
 1270 1280 ...
 ... ALA THR LEU ILE GLY SER ASP PHE ASN ASP HIS
 ... GCCACTCTTATCGGTTTCGGATTTTAACGATCA
 ...1290 1300 1310 1320

173/235

LYS LYS PRO LEU ASN ILE LYS GLY ASP ...
 TAAACAACCTTTAAATAAAGGAGAT...
 1330 1340 ...
 ... VAL VAL ASN ARG GLY ASN LEU THR ALA GLY GLY
 ... GTCGTCAATAGAGGCACCTTACCCTGGCGG
 ...1350 1360 1370 1380

ASN VAL ILE ASN ILE GLY GLY ASN LEU ...
 CAATGTTATCAATAATAGCGGGAATCTT...
 1390 1400 ...
 ... THR VAL GLU ASN GLY ALA ASN LEU LYS ALA ILE
 ... ACCGTTGAAATAAGCGCCCAATCTTAAGCTAT
 ...1410 1420 1430 1440

174/235

FIG.27H

```

THR ASN PHE THR PHE ASN VAL GLY GLY ...
C A C A A A T T T C A C T T T T A A T G T A G G C G G C ...
1450
...
... LEU PHE ASN ASN LYS GLY ASN SER ASN ILE SER
... T T G T T T A A C A A C A A G G C A A T T C A A A T A T C T C
1470
...1480
1490
1500

```

```

ILE ALA ARG GLY GLY ALA LYS PHE LYS ...
C A T G C T A G A G G A G G G C T A A A T T T A A A ...
1510
...
... ASP ILE ASN ASN THR SER SER LEU ASN ILE THR
... G A T A T C A A T A A C A C C A G T A G C T T A A A T A T T A C
1530
...1540
1550
1560

```

```

THR ASN SER ASP THR THR TYR ARG THR ...
C A C C A A C T C C G A C A C C A C T T A C C G T A C C ...
1570
...
... ILE ILE GLU GLY ASN ILE THR ASN LYS ALA GLY
... A T T A T A G A A G G T A A T A T A C C A A C A A G C A G G
1590
...1600
1610
1620

```

```

ASP LEU ASN ILE ILE ASP ASN LYS GLY ...
T G A T T G A A T A T C A T T G A T A A T A A A G G T ...
1630
1640
...

```

FIG.271

SUBSTITUTE SHEET (RULE 26)

... ASN ALA GLU ILE GLN ILE GLY GLY ASN ILE SER
 ... AACGCTGAATAATCGCAATGGCGGCAACATCTC
 ...1650 1660 1670 1680

GLN LYS GLU GLY ASN LEU THR ILE SER ...
 GCAAAAGAGGTAACTCAGATTTC...
 ...1690 1700 ...
 ... SER ASP LYS ILE ASN ILE THR LYS GLN ILE THR
 ... TCCGATAAATACTCAATATTACCAACAGATAAC
 ...1710 1720 1730 1740

ILE LYS LYS GLY VAL ASN GLY GLU ASN ...
 ATCAAGAGGGTGTTAACGGAGAGAAC...
 ...1750 1760 ...
 ... SER ASP SER SER THR LYS SER SER GLN ALA ASN LEU
 ... TCTGATTCAAGTACGAAAGTCAAGCCAAATCT
 ...1770 1780 1790 1800

THR ILE LYS LYS THR LYS GLU LEU LYS LEU ...
 ACCATTAAACCAAGAAATTGAAATTA...
 ...1810 1820 ...
 ... THR GLN ASP LEU ASN ILE SER GLY PHE ASN LYS
 ... ACACAGACCTAAATAATTTCAGGCTTCAACA
 ...1830 1840 1850 1860

175 / 235

176/235

FIG.27J

ALA LYS ILE VAL ALA LYS ASP SER SER ...
 AGCAAAGATTGTAGCTAAAGATAGTAGT...
 1870 1880
 ... ASN LEU THR ILE GLY ASN SER ASP SER GLY
 ... AATTAACTATTGGTAAATAGTAGTAGCGG
 ...1890 1900 1910 1920

ASN THR SER ALA LYS THR VAL THR PHE ...
 CAATACTAGCGCTAAACAGTAACTTT...
 1930 1940
 ... ASN VAL LYS ASP SER LYS ILE SER ALA ASP
 ... AACAAATGTTAAAGATTCAAAATCTCTGCTGA
 ...1950 1960 1970 1980

GLY HIS LYS VAL THR LEU ASN SER LYS ...
 CGGTCAACAAGGTGACACTAAATAGCAA...
 1990 2000
 ... VAL LYS THR LEU SER ASP ASN ASP ASN THR
 ... GTGAAACACTTAGTGATAATGATAACAAC
 ...2010 2020 2030 2040

GLU GLY GLY SER ASP ASN THR GLY ...
 TGAAGGTGGCAGTGACACAAATACCGGT...
 2050 2060 ...

09/806709

FIG.27K

SUBSTITUTE SHEET (RULE 26)

... LEU THR ILE THR ALA LYS ASP VAL GLU VAL ASN
 ... TTA A C T A T T A C T G C A A A G A T G T A G A G T A A A
 ...2070 2080 2090 2100

ASN ASN ILE THR SER HIS LYS THR VAL ...
 C A A C A A T A T T A C T T C T C A C A A A C A G T G ...
 2110 2120

... ASN VAL SER ALA ALA ASN GLY GLY ILE THR THR
 ... A A C G T C T C T G C G G C A A A T G G A G G G A T T A C C A C
 ...2130 2140 2150 2160

LYS THR GLY THR THR ILE ASN ALA THR ...
 T A A A C A G G T A C A C C A T T A A T G C A A C C ...
 2170 2180

... ALA GLY ASN VAL GLU ILE THR ALA HIS THR GLY
 ... G C C G G T A A C G T G G A G A T A A C C G C T C A T A C A G G
 ...2190 2200 2210 2220

SER ILE GLN GLY GLY ILE GLU SER LYS ...
 C A G T A T C C A A G G C G G A A T T G A G T C C A A G ...
 2230 2240

... PRO GLY SER VAL THR ILE VAL ALA GLY GLY ASP
 ... C C T G G C T C T G T G A C A A T T G T G G C A G G C G G C G A
 ...2250 2260 2270 2280

178/ 235

FIG.27L

THR LEU ALA VAL GLY ASN ILE SER GLY ...
 T A C T C T T G C T G T A G G T A A T A T T T C A G G C ...
 2290 2300
 ... ASN ALA VAL THR VAL THR ALA ASN SER GLY ALA
 ... A A C G C C G T T A C T G T T A C T G C C A A A T A G C G G T G C
 ...2310 2320 2330 2340

LEU THR THR LEU ALA GLY SER THR ILE ...
 A T T A C C A C T T T G G C A G G C T C T A C A A T T ...
 2350 2360
 ... LYS GLY THR GLU SER ILE THR SER SER GLN
 ... A A G G A A C C G A G A G T A T A C C A C T T C A A G T C A
 ...2370 2380 2390 2400

SER GLY ASN ILE GLY GLY LYS ILE SER ...
 A T C A G G T A A T A T C G G C G G T A A A T T T C C ...
 2410 2420
 ... GLY LYS THR VAL ASN VAL LYS ALA THR ASN SER
 ... G G C A A G A C A G T A A A C G T T A A A G C A A C T A A T A G
 ...2430 2440 2450 2460

LEU THR THR GLN ALA ASP SER LYS ILE ...
 T T T A C C A C C C A A G C A G A C T C A A A A T T ...
 2470 2480
 ...

FIG.27M

179 / 235

... GLU ALA THR GLU GLY GLU ALA ASN VAL THR SER
 ... GAAGCGACTGAAGCGCGAGCTAATGTAAACAAG
 ...2490 2500 2510 2520

LYS THR SER ILE ILE GLY GLY THR ILE ...
 CAAACAAGCATATTTGGCGGTACAAATT...
 2530 2540

... SER GLY GLY THR VAL GLU VAL THR ALA THR GLU
 ... TCTGGTGCCACAGTAGAAGTTACCGCGACCGA
 ...2550 2560 2570 2580

GLY LEU THR THR GLN ALA GLY SER THR ...
 AGGTTTAACCAACCAAGCAGGCTCTACG...
 2590 2600
 ... ILE THR GLY THR GLU SER VAL THR THR SER SER
 ... ATTA CTGGAACCGAGAGCGTGACCACTTCAAG
 ...2610 2620 2630 2640

GLN SER GLY ASN ILE GLY GLY MET ILE ...
 CCATCAGGTATAATCGGCGGCATGATT...
 2650 2660
 ... SER GLY GLY LYS VAL GLU VAL SER ALA THR LYS
 ... TCTGGTGCCAAAGTAGAAGTTAGCGCAACCA
 ...2670 2680 2690 2700

FIG.27N

```

ASP  LEU  ILE  THR  LYS  SER  GLY  SER  GLU  ...
AGATTAAATTACTAATAATCCGGTTTCAGAG...
2710
... ILE  LYS  ALA  THR  ALA  GLY  GLU  VAL  ASN  VAL  THR
... ATTAAAGCAACGGCGGCGAGGTGAATGTAAAC
...2730
2740
2750
2760

```

180/235

```

SER  ALA  THR  GLY  THR  ILE  ASP  GLY  THR  ...
AGTGCAACAGGTACAAATTGACGGTACG...
2770
... ILE  SER  GLY  ASN  THR  VAL  ASN  VAL  THR  ALA  ASN
... ATTCCGGTAATAACGGTAATAATGTACAGCAAA
...2790
2800
2810
2820

```

```

THR  GLY  ASP  LEU  THR  VAL  GLU  ASP  ALA  ...
TAC TGGCGATTTAAC TGT TGAAGATGCC...
2830
... ALA  LYS  ILE  ASP  ALA  THR  GLY  ALA  ALA  THR
... GCAAAATTGATGCCGACAGGAGGAGCCGGAC
...2850
2860
2870
2880

```

```

LEU  THR  ALA  THR  SER  GLY  LYS  LEU  THR  ...
CCTAAC TGCAACA TC GGGCAATA TTAACC...
2890
2900
...

```

FIG.270

```

...   THR   LYS   ALA   SER   SER   SER   ILE   THR   SER   ALA   ASN
...   A C T A A G G C T A G T T C A A G C A T T A C T T C A G C T A A
...2910                                     2930                2940

ASN   GLN   VAL   ASN   LEU   SER   ALA   LYS   ASP ...
T A A C C A G G T A A A C C T T T C A G C T A A G G A T ...
2950
...   GLY   SER   ILE   GLY   GLY   ASN   ILE   ASN   ALA   ALA   ASN
...   G G T A G C A T T G G G G G A A T A T C A A T G C T G C T A A
...2970                                     2990                3000
181/235

VAL   THR   LEU   ASN   THR   THR   GLY   ALA   LEU ...
T G T A A C A C T G A A T A C T A C A G G C G C T C T A ...
3010
...   THR   THR   VAL   LYS   GLY   SER   SER   ILE   ASN   ALA   ASN
...   A C T A C C G T G A A G G G T T C A A G C A T T A A C G C A A A
...3030                                     3050                3060

SER   GLY   THR   LEU   VAL   ILE   ASN   ALA   LYS ...
C A G C G G C A C C T T G G T T A T T A A C G C A A A ...
3070
...   ASP   ALA   GLU   LEU   ASN   GLY   GLU   ALA   SER   GLY   ASN
...   G A C G C T G A G C T A A A T G G T G A G G C A T C A G G T A A
...3090                                     3110                3120

```

FIG.27P

HIS THR VAL VAL ASN ALA THR ASN ALA ...
 CCATACAGTAGTGAAATGCCAACCAACGCA...
 3130 3140
 ... ASN GLY SER GLY SER VAL ILE ALA THR THR SER
 ... AATGGCTCCGGCAGCGTAATCGCGACACCTC
 ...3150 3160 3170 3180
 182/235
 SER ARG VAL ASN ILE THR GLY ASP LEU ...
 AGCAGAGTGAAACATCATCTGGGGATTTA...
 3190 3200
 ... ILE THR ILE ASN GLY LEU ASN ILE ILE SER LYS
 ... ATCAACAATAAATGGATTAAATAATCTTCAA
 ...3210 3220 3230 3240
 ASN GLY ILE ASN THR VAL LEU LEU LYS ...
 AACGGTATAAACACCGTACTGTTAATAA...
 3250 3260
 ... GLY VAL LYS ILE ASP VAL LYS TYR ILE GLN PRO
 ... GCGTTAAATAATTGATGTGAATAATCTCAACC
 ...3270 3280 3290 3300
 GLY ILE ALA SER VAL ASP GLU VAL ILE ...
 GGGTATAGCAAGCGTAGTAGTAATT...
 3310 3320 ...

FIG.27Q

```

...   GLU   ALA   LYS   ARG   ILE   LEU   GLU   LYS   VAL   LYS   ASP
...   G A G C G A A A C G C A T C C T T G A G A A G G T A A A G A
...3330                                     3350       3360

LEU   SER   ASP   GLU   GLU   ARG   GLU   ALA   LEU ...
T T T A T C T G A T G A A G A A G A G C G T T A...
3370                                     3380
...   ALA   LYS   LEU   GLY   VAL   SER   ALA   VAL   ARG   PHE   ALA
...   G C T A A A C T T G G C G T A A G C G C T G T A C G T T T T G C
...3390                                     3410       3420

GLU   PRO   ASN   ASN   ALA   ILE   THR   ILE   ASN ...
T G A G C C A A A T A A T G C C A T T A C G A T T A A T...
3430                                     3440
...   THR   GLN   ASN   GLU   PHE   THR   THR   ARG   PRO   SER   SER
...   A C A C A A A A T G A G T T T A C A C C A G A C C A T C A A G
...3450                                     3470       3480

GLN   VAL   THR   ILE   SER   GLU   GLY   LYS   VAL ...
T C A A G T G A C A A T T T C T G A A G G T A A G G T A...
3490                                     3500
...   CYS   PHE   LEU   ILE   GLY   ASN   GLY   ALA   THR   ILE   CYS
...   T G T T T C T T A A T C G G C A A T G G T G C A A C A A T A T G
...3510                                     3530       3540

```

184/235

FIG.27R

THR	ASN	ILE	ALA	ASP	ILE	GLU	ARG	***
C	A	C	C	A	A	T	A	T
T	G	C	T	G	A	T	A	T
G	A	T	T	G	A	T	G	A
G	G	T	A	G	C	C	G	G
T	A	G	C	C	G	G	T	A
A	G	T	A	G	C	C	G	T
								A

3550 3560

SUBSTITUTE SHEET (RULE 26)

FIG. 28B

TTT TTA CAA GAA AAC AAC AAC TCC GCC GGA TTC AAC CGT GTT ACA TCT 740
Phe Leu Gln Glu Asn Asn Asn Ser Ala Val Phe Asn Arg Val Thr Ser 130
115 120 125

187/235

FIG.28C

AAC CAA ATC TCC CAA TTA AAA GGG ATT TTA GAT TCT AAC GGA CAA GTC	788
Asn Gln Ile Ser Gln Leu Lys Gly Ile Thr Ile Gly Lys Asp Ser Asn Gly Gln Val	145
	135
	140
TTT TTA ATC AAC CCA AAT GGT ATC ACA ATA GGT AAA GAC GCA ATT ATT	836
Phe Leu Ile Asn Pro Asn Gly Ile Thr Ile Gly Lys Asp Ala Ile Ile	160
	150
	155
AAC ACT AAT GGC TTT ACG GCT TCT ACG CTA GAC ATT TCT AAC GAA AAC	884
Asn Thr Asn Gly Phe Thr Ala Ser Thr Leu Asp Ile Ser Asn Glu Asn	175
	165
	170
ATC AAG GCG CGT AAT TTC ACC TTC GAG CAA ACC AAA GAT AAA GCG CTC	932
Ile Lys Ala Arg Asn Phe Thr Phe Glu Gln Thr Lys Asp Lys Ala Leu	190
	180
	185
GCT GAA ATT GTG AAT CAC GGT TTA ATT ACT GTC GGT AAA GAC GGC AGT	980
Ala Glu Ile Val Asn His Gly Leu Ile Thr Val Gly Lys Asp Gly Ser	210
	195
	200
	205
GTA AAT CTT ATT GGT GGC AAA GTG AAA AAC GAG GGT GIG ATT AGC GTA	1028
Val Asn Leu Ile Gly Gly Lys Val Lys Asn Glu Gly Val Ile Ser Val	225
	215
	220

188/235

FIG.28D

AAT GGT GGC AGC ATT TCT TTA CTC GCA GGG CAA AAA ATC ACC ATC AGC	1076
Asn Gly Gly Ser Ile Ser Leu Leu Ala Gly Gln Lys Ile Thr Ile Ser	
230 235 240	
GAT ATA ATA AAC CCA ACC ATT ACT TAC AGC ATT GCC GCG OCT GAA AAT	1124
Asp Ile Ile Asn Pro Thr Ile Thr Tyr Ser Ile Ala Ala Pro Glu Asn	
245 250 255	
GAA GCG GTC AAT CTG GGC GAT ATT TTT GCC AAA GGC GGT AAC ATT AAT	1172
Glu Ala Val Asn Leu Gly Asp Ile Phe Ala Lys Gly Gly Asn Ile Asn	
260 265 270	
GTC CGT GCT GCC ACT ATT CGA AAC CAA GGT AAA CTT TCT GCT GAT TCT	1220
Val Arg Ala Ala Thr Ile Arg Asn Gln Gly Lys Leu Ser Ala Asp Ser	
275 280 285 290	
GTA AGC AAA GAT AAA AGC AGC AAT ATT GGT CTT TOC GCC AAA GAG GGT	1268
Val Ser Lys Asp Lys Ser Gly Ser Gly Asn Ile Val Leu Ser Ala Lys Glu Gly	
295 300 305	
GAA GCG GAA ATT GCG GGT GFA ATT TOC GCT CAA AAT CAG CAA GCT AAA	1316
Glu Ala Glu Ile Gly Gly Val Ile Ser Ala Gln Asn Gln Gln Ala Lys	
310 315 320	

189/235

FIG.28E

GGC GGC AAG CTG ATG ATT ACA GGC GAT AAA GTC ACA TTA AAA ACA GGT Gly Gly Lys Leu Met Ile Thr Gly Asp Lys Val Thr Leu Lys Thr Gly 325 330 335	1364
GCA GTT ATC GAC CTT TCA GGT AAA GAA GGC GGA ACT TAC CTT GGC Ala Val Ile Asp Leu Ser Gly Lys Glu Gly Gly Thr Tyr Leu Gly 340 345 350	1412
GGT GAC GAG CGC GGC GAA GGT AAA AAG GGC ATT CAA TTA GCA AAG AAA Gly Asp Glu Arg Gly Glu Gly Lys Lys Gly Ile Gln Leu Ala Lys Lys 355 360 365 370	1460
ACC TCT TTA GAA AAA GGC TCA ACC ATC AAT GTA TCA GGC AAA GAA AAA Thr Ser Leu Glu Lys Gly Ser Thr Ile Asn Val Ser Gly Lys Glu Lys 375 380 385	1508
GGC GGA CGC GCT ATT GTG TGG GGC GAT ATT GCG TTA ATT GAC GGC AAT Gly Gly Arg Ala Ile Val Trp Gly Asp Ile Ala Leu Ile Asp Gly Asn 390 395 400	1556
ATT AAC GCT CAA GGT AGT GGT GAT ATC GCT AAA ACC GGT TTT GTG Ile Asn Ala Gln Gly Ser Gly Asp Ile Ala Lys Thr Gly Gly Phe Val 405 410 415	1604

09/806709

190/235

FIG.28F

GAG ACG TCG GCG CAT GAT TTA TTC ATC AAA GAC AAT GCA ATT GTT GAC Glu Thr Ser Gly His Asp Leu Phe Ile Lys Asp Asn Ala Ile Val Asp	1652
420 425 430	
↓	
GCC AAA GAG TGG TTG TTA GAC CCG GAT AAT GTA TCT ATT AAT GCA GAA Ala Lys Glu Trp Leu Leu Asp Pro Asp Asn Val Ser Ile Asn Ala Glu	1700
435 440 445 450	
ACA GCA GGA CGC AGC AAT ACT TCA GAA GAC GAT GAA TAC ACG GGA TCC Thr Ala Gly Arg Ser Asn Thr Ser Glu Asp Asp Glu Tyr Thr Gly Ser	1748
455 460 465	
GCG AAT AGT GCC AGC ACC CCA AAA CGA AAC AAA GAA AAG ACA ACA TTA Gly Asn Ser Ala Ser Thr Pro Lys Arg Asn Lys Lys Glu Lys Thr Thr Leu	1796
470 475 480	
ACA AAC ACA ACT CTT GAG AGT ATA CTA AAA AAA GGT ACC TTT GTT AAC Thr Asn Thr Thr Leu Glu Ser Ile Leu Lys Lys Gly Thr Phe Val Asn	1844
485 490 495	
ATC ACT GCT AAT CAA CGC ATC TAT GTC AAT AGC TCC ATT AAT TTA TCC Ile Thr Ala Asn Gln Arg Ile Tyr Val Asn Ser Ser Ile Asn Leu Ser	1892
500 505 510	

191/235

FIG.28G

AAT GGC AGC TTA ACT CTT TGG AGT GAG GGT CCG AGC GGT GGC GGC GTT Asn Gly Ser Leu Thr Leu Trp Ser Glu Gly Arg Ser Gly Gly Gly Val 515 520 525 530	1940
GAG ATT AAC AAC GAT ATT ACC ACC GGT GAT GAT ACC AGA GGT GCA AAC Glu Ile Asn Asn Asp Ile Thr Thr Gly Asp Asp Thr Arg Gly Ala Asn 535 540 545	1988
TTA ACA ATT TAC TCA GGC GGC TGG GTT GAT GTT CAT AAA AAT ATC TCA Leu Thr Ile Tyr Ser Gly Gly Trp Val Asp Val His Lys Asn Ile Ser 550 555 560	2036
CTC GGC GCG CAA GGT AAC ATA AAC ATT ACA GCT AAA CAA GAT ATC GCC Leu Gly Ala Gln Gly Asn Ile Asn Ile Thr Ala Lys Gln Asp Ile Ala 565 570 575	2084
TTT GAG AAA GGA AGC AAC CAA GTC ATT ACA GGT CAA GGC ACT ATT ACC Phe Glu Lys Gly Ser Asn Gln Val Ile Thr Gly Gln Gly Thr Ile Thr 580 585 590	2132
TCA GGC AAT CAA AAA GGT TTT AGA TTT AAT AAT GTC TCT CTA AAC GGC Ser Gly Asn Gln Lys Gly Phe Arg Phe Asn Asn Val Ser Leu Asn Gly 595 600 605 610	2180

192 / 235

FIG.28H

ACT GGC AGC GGA CTG CAA TTC ACC ACT AAA AGA ACC AAT AAA TAC GCT Thr Gly Ser Gly Leu Gln Phe Thr Thr Lys Arg Thr Asn Lys Tyr Ala 615 620 625	2228
ATC ACA AAT AAA TTT GAA GGG ACT TTA AAT ATT TCA GGG AAA GIG AAC Ile Thr Asn Lys Phe Glu Gly Thr Leu Asn Ile Ser Gly Lys Val Asn 630 635 640	2276
ATC TCA ATG GTT TTA CCT AAA AAT GAA AGT GGA TAT GAT AAA TTC AAA Ile Ser Met Val Leu Pro Lys Asn Glu Ser Gly Tyr Asp Lys Phe Lys 645 650 655	2324
GGA CGC ACT TAC TGG AAT TTA ACC TCC TTA AAT GTT TCC GAG AGT GGC Gly Arg Thr Tyr Trp Asn Leu Thr Ser Leu Asn Val Ser Glu Ser Gly 660 665 670	2372
GAG TTT AAC CTC ACT ATT GAC TCC AGA GGA AGC GAT AGT GCA GGC ACA Glu Phe Asn Leu Thr Ile Asp Ser Arg Gly Ser Asp Ser Ala Gly Thr 675 680 685 690	2420
CTT ACC CAG CCT TAT AAT TTA AAC GGT ATA TCA TTC AAC AAA GAC ACT Leu Thr Gln Pro Tyr Asn Leu Asn Gly Ile Ser Phe Asn Lys Asp Thr 695 700 705	2468

193/235

FIG.28I

ACC TTT AAT GTT GAA CGA AAT GCA AGA GTC AAC TTT GAC ATC AAG GCA	2516
Thr Phe Asn Val Glu Arg Asn Ala Arg Val Asn Phe Asp Ile Lys Ala	
710 715 720	
CCA ATA GGG ATA AAT AAG TAT TCT AGT TTG AAT TAC GCA TCA TTT AAT	2564
Pro Ile Gly Ile Asn Lys Tyr Ser Ser Leu Asn Tyr Ala Ser Phe Asn	
725 730 735	
GGA AAC ATT TCA GTT TCG GGA GGG GGG AGT GTT GAT TTC ACA CTT CTC	2612
Gly Asn Ile Ser Val Ser Gly Gly Gly Ser Val Asp Phe Thr Leu Leu	
740 745 750	
GCC TCA TCC TCT AAC GTC CAA ACC CCC GGT GTA GTT ATA AAT TCT AAA	2660
Ala Ser Ser Ser Asn Val Val Gln Thr Pro Gly Val Val Ile Asn Ser Lys	
755 760 765 770	
TAC TTT AAT GTT TCA ACA GGG TCA AGT TTA AGA TTT AAA ACT TCA GGC	2708
Tyr Phe Asn Val Ser Thr Gly Ser Ser Leu Arg Phe Lys Thr Ser Gly	
775 780 785	
TCA ACA AAA ACT GGC TTC TCA ATA GAG AAA GAT TTA ACT TTA AAT GCC	2756
Ser Thr Lys Thr Gly Phe Ser Ile Glu Lys Asp Leu Thr Leu Asn Ala	
790 795 800	

194/235

FIG.28J

ACC GGA GGC AAC ATA ACA CTT TTG CAA GTT GAA GGC ACC GAT GGA ATG Thr Gly Gly Asn Ile Thr Leu Leu Gln Val Glu Gly Thr Asp Gly Met 805 810 815	2804
ATT GGT AAA GGC ATT GTA GCC AAA AAC ATA ACC TTT GAA GGA GGT Ile Gly Lys Gly Ile Val Ala Lys Lys Asn Ile Thr Phe Glu Gly Gly 820 825 830	2852
AAC ATC ACC TTT GGC TCC AGG AAA GCC GTA ACA GAA ATC GAA GGC AAT Asn Ile Thr Phe Gly Ser Arg Lys Ala Val Thr Glu Ile Glu Gly Asn 835 840 845 850	2900
GTT ACT ATC AAT AAC AAC GCT AAC GTC ACT CTT ATC GGT TCG GAT TTT Val Thr Ile Asn Asn Ala Asn Val Thr Leu Ile Gly Ser Asp Phe 855 860 865	2948
GAC AAC CAT CAA AAA CCT TTA ACT ATT AAA AAA GAT GTC ATC ATT AAT Asp Asn His Gln Lys Pro Leu Thr Ile Lys Lys Asp Val Ile Ile Asn 870 875 880	2996
AGC GGC AAC CTT ACC GCT GGA GGC AAT ATT GTC AAT ATA GCC GGA AAT Ser Gly Asn Leu Thr Ala Gly Gly Asn Ile Val Asn Ile Ala Gly Asn 885 890 895	3044

195/235

FIG.28K

CTT ACC GGT GAA AGT AAC GCT AAT TTC AAA GCT ATC ACA AAT TTC ACT	3092
Leu Thr Val Glu Ser Asn Ala Asn Phe Lys Ala Ile Thr Asn Phe Thr	
900	905
	910
TTT AAT GTA GGC GGC TTG TTT GAC AAC AAA GGC AAT TCA AAT ATT TCC	3140
Phe Asn Val Gly Gly Leu Phe Asp Asn Lys Gly Asn Ser Asn Ile Ser	
915	920
	925
ATT GCC AAA GGA GGG GCT GGC TTT AAA GAC AAT GAT AAT TCC AAG AAT	3188
Ile Ala Lys Gly Gly Ala Arg Phe Lys Asp Ile Asp Asn Ser Lys Asn	
	935
	940
	945
TTA AGC ATC ACC ACC AAC TCC AGC TCC ACT TAC CGC ACT ATT ATA AGC	3236
Leu Ser Ile Thr Thr Asn Ser Ser Ser Thr Tyr Arg Thr Ile Ile Ser	
	950
	955
	960
GGC AAT ATA ACC AAT AAA AAC GGT GAT TTA AAT ATT ACG AAC GAA GGT	3284
Gly Asn Ile Thr Asn Lys Asn Gly Asp Leu Asn Ile Thr Asn Glu Gly	
	965
	970
	975
AGT GAT ACT GAA ATG CAA ATT GGC GGC GAT GTC TCG CAA AAA GAA GGT	3332
Ser Asp Thr Glu Met Gln Ile Gly Gly Asp Val Ser Gln Lys Glu Gly	
	980
	985
	990

196/235

FIG.28L

AAT CTC ACG ATT TCT TCT GAC AAA ATC AAT ATT ACC AAA CAG ATA ACA	3380
Asn Leu Thr Ile Ser Ser Asp Lys Ile Asn Ile Thr Lys Gln Ile Thr	
995 1000 1005 1010	
ATC AAG GCA GGT GTT GAT GGG GAG AAT TCC GAT TCA GAC GCG ACA AAC	3428
Ile Lys Ala Gly Val Asp Gly Glu Asn Ser Asp Ser Asp Ala Thr Asn	
1015 1020 1025	
AAT GCC AAT CTA ACC ATT AAA ACC AAA GAA TTG AAA TTA ACG CAA GAC	3476
Asn Ala Asn Leu Thr Ile Lys Thr Lys Glu Leu Lys Leu Thr Gln Asp	
1030 1035 1040	
CTA AAT ATT TCA GGT TTC AAT AAA GCA GAG ATT ACA GCT AAA GAT GGT	3524
Leu Asn Ile Ser Gly Phe Asn Lys Ala Glu Ile Thr Ala Lys Asp Gly	
1045 1050 1055	
AGT GAT TTA ACT ATT GGT AAC ACC AAT AGT GCT GAT GGT ACT AAT GCC	3572
Ser Asp Leu Thr Ile Gly Asn Thr Asn Ser Ala Asp Gly Thr Asn Ala	
1060 1065 1070	
AAA AAA GTA ACC TTT AAC CAG GTT AAA GAT TCA AAA ATC TCT GCT GAC	3620
Lys Lys Val Thr Phe Asn Gln Val Lys Asp Ser Lys Ile Ser Ala Asp	
1075 1080 1085 1090	

197/235

FIG.28M

GGT CAC AAG GTG ACA CTA CAC AGC AAA GTG GAA ACA TCC GGT AGT AAT Gly His Lys Val Thr Leu Thr His Ser Lys Val Glu Thr Ser Gly Ser Asn 1095 1100 1105	3668
AAC AAC ACT GAA GAT AGC AGT GAC AAT AAT GCC GGC TTA ACT ATC GAT Asn Asn Thr Glu Asp Ser Ser Asp Asn Asn Ala Gly Leu Thr Ile Asp 1110 1115 1120	3716
GCA AAA AAT GTA ACA GTA AAC AAC AAT ATT ACT TCT CAC AAA GCA GTG Ala Lys Asn Val Thr Val Asn Asn Ile Thr Ser His Lys Ala Val 1125 1130 1135	3764
AGC ATC TCT GCG ACA AGT GGA GAA ATT ACC ACT AAA ACA GGT ACA ACC Ser Ile Ser Ala Thr Ser Gly Glu Ile Thr Thr Lys Thr Gly Thr Thr 1140 1145 1150	3812
ATT AAC GCA ACC ACT GGT AAC GTG GAG ATA ACC GGT CAA ACA GGT AGT Ile Asn Ala Thr Thr Gly Asn Val Glu Ile Thr Ala Gln Thr Gly Ser 1155 1160 1165 1170	3860
ATC CTA GGT GGA ATT GAG TCC AGC TCT GGC TCT GTA ACA CTT ACT GCA Ile Leu Gly Gly Ile Glu Ser Ser Gly Ser Val Thr Leu Thr Ala 1175 1180 1185	3908

198/235

FIG.28N

ACC GAG GGC GCT CTT GCT GTA AGC AAT ATT TCG GGC AAC ACC GTT ACT	3956
Thr Glu Gly Ala Leu Ala Val Ser Asn Ile Ser Gly Asn Thr Val Thr	
1190 1195 1200	
GTT ACT GCA AAT AGC GGT GCA TTA ACC ACT TTG GCA GGC TCT ACA ATT	4004
Val Thr Ala Asn Ser Gly Ala Leu Thr Thr Leu Ala Gly Ser Thr Ile	
1205 1210 1215	
AAA GGA ACC GAG AGT GTA ACC ACT TCA AGT CAA TCA GGC GAT ATC GGC	4052
Lys Gly Thr Glu Ser Val Thr Thr Ser Ser Gln Ser Gly Asp Ile Gly	
1220 1225 1230	
GGT ACG ATT TCT GGT GGC ACA GTA GAG GTT AAA GCA ACC GAA AGT TTA	4100
Gly Thr Ile Ser Gly Gly Thr Val Glu Val Lys Ala Thr Glu Ser Leu	
1235 1240 1245 1250	
ACC ACT CAA TCC AAT TCA AAA ATT AAA GCA ACA ACA GGC GAG GCT AAC	4148
Thr Thr Gln Ser Ser Asn Ser Lys Ile Lys Ala Thr Thr Gly Glu Ala Asn	
1255 1260 1265	
GTA ACA AGT GCA ACA GGT ACA ATT GGT GGT ACG ATT TCC GGT AAT ACG	4196
Val Thr Ser Ala Thr Gly Thr Ile Gly Gly Thr Ile Ser Gly Asn Thr	
1270 1275 1280	

199/235

FIG.280

GTA AAT GTT ACG GCA AAC GCT GGC GAT TTA ACA GTT GGG AAT GGC GCA Val Asn Val Thr Ala Asn Ala Gly Asp Leu Thr Val Gly Asn Gly Ala 1285 1290 1295	4244
GAA ATT AAT GCG ACA GAA GGA GCT GCA ACC TTA ACT ACA TCA TCG GGC Glu Ile Asn Ala Thr Glu Gly Ala Ala Thr Leu Thr Thr Ser Ser Gly 1300 1305 1310	4292
AAA TTA ACT ACC GAA GCT AGT TCA CAC ATT ACT TCA GCC AAG GGT CAG Lys Leu Thr Thr Glu Ala Ser Ser His Ile Thr Ser Ala Lys Gly Gln 1315 1320 1325 1330	4340
GTA AAT CTT TCA GCT CAG GAT GGT AGC GTT GCA GGA AGT ATT AAT GGC Val Asn Leu Ser Ala Gln Asp Gly Ser Val Ala Gly Ser Ile Asn Ala 1335 1340 1345	4388
GCC AAT GTG ACA CTA AAT ACT ACA GGC ACT TTA ACT ACC GTG AAG GGT Ala Asn Val Thr Leu Asn Thr Thr Gly Thr Leu Thr Thr Val Lys Gly 1350 1355 1360	4436
TCA AAC ATT AAT GCA ACC AGC GGT ACC TTG GTT ATT AAC GCA AAA GAC Ser Asn Ile Asn Ala Thr Ser Gly Thr Leu Val Ile Asn Ala Lys Asp 1365 1370 1375	4484

200/235

FIG.28P

GCT GAG CTA AAT GGC GCA GCA TTG GGT AAC CAC ACA GTG GTA AAT GCA Ala Glu Leu Asn Gly Ala Ala Leu Gly Asn His Thr Val Val Asn Ala 1380 1385 1390	4532
ACC AAC GCA AAT GGC TCC GGC AGC GTA ATC GCG ACA ACC TCA AGC AGA Thr Asn Ala Asn Gly Ser Gly Ser Val Ile Ala Thr Thr Ser Ser Arg 1395 1400 1405 1410	4580
GTG AAC ATC ACT GGG GAT TTA ATC ACA ATA AAT GGA TTA AAT ATC ATT Val Asn Ile Thr Gly Asp Leu Ile Thr Ile Asn Gly Leu Asn Ile Ile 1415 1420 1425	4628
TCA AAA AAC GGT ATA AAC ACC GTA CTG TTA AAA GGC GTT AAA ATT GAT Ser Lys Asn Gly Ile Asn Thr Val Leu Leu Lys Gly Val Lys Ile Asp 1430 1435 1440	4676
GTG AAA TAC ATT CAA CCG GGT ATA GCA AGC GTA GAT GAA GTA ATT GAA Val Lys Tyr Ile Gln Pro Gly Ile Ala Ser Val Asp Glu Val Ile Glu 1445 1450 1455	4724
GGG AAA CCG ATC CTT GAG AAG GTA AAA GAT TTA TCT GAT GAA GAA AGA Ala Lys Arg Ile Leu Glu Lys Val Lys Asp Leu Ser Asp Glu Glu Arg 1460 1465 1470	4772

09/806709

201/235

FIG.28Q

GAA GCG TTA GCT AAA CTT GGA GTA AGT GCT GTA CGT TTT ATT GAG CCA Glu Ala Leu Ala Lys Leu Gly Val Ser Ala Val Arg Phe Ile Glu Pro 1475 1480 1485 1490	4820
AAT AAT ACA ATT ACA GTC GAT ACA CAA AAT GAA TTT GCA ACC AGA CCA Asn Asn Thr Ile Thr Val Asp Thr Gln Asn Glu Phe Ala Thr Arg Pro 1495 1500 1505	4868
TTA AGT CGA ATA GIG ATT TCT GAA GGC AGG GCG TGT TTC TCA AAC AGT Leu Ser Arg Ile Val Ile Ser Glu Gly Arg Ala Cys Phe Ser Asn Ser 1510 1515 1520	4916
GAT GGC GCG ACG GIG TGC GTT AAT ATC GCT GAT AAC GCG CCG Asp Gly Ala Thr Val Cys Val Asn Ile Ala Asp Asn Gly Arg 1525 1530 1535	4958
TACGGGTCAG TAATTGACAA GGTAGATTTC ATCCGCAAT GAAGTCATTT TATTTTCGTA	5018
TTATTTACIG TGIGGGTTAA AGTTCAGTAC GGGCTTTACC CATCTTGTAA AAAATTACGG	5078
AGAATACAAT AAAGTATTTT TAACAGGTAA TTATTATG	5116

202/235

FIG.29A

TAAATATACA AGATAATATAA AATAAATCAA GATTTTIGIG ATGACAAACA ACAATTACAA	60
CACCTTTTTT GCAGTCTATA TGCAATATAT TTAAAAAAT AGTATAAATC CGCATATATA	120
AATGGTATAA TCCTTTCATCT TTCAATCTTAA ATCTTTCATC TTTCATCTTT CATCTTTCAT	180
CTTTCATCTT TCATCTTICA TCCTTTCATCT TTTCATCTTTC ATCTTTCATC TTTCATCTTT	240
CACATGAAAT GATGAACCGA CGGAAGCGAG CGAGGGGCAA GAATGAAGAG GGAGCTGAAC	300
GAACGCCAAT GATAAAGTAA TTTAATTGTT CAACTAACCT TAGGAGAAAA T ATG AAC	357
	Met Asn 1
AAG ATA TAT CGT CTC AAA TTC AGC AAA CGC CTG AAT GCT TTG GTT GCT	405
Lys Ile Tyr Arg Leu Lys Phe Ser Lys Arg Leu Asn Ala Leu Val Ala	15
	5 10
GTG TCT GAA TTG GCA CGG GGT TGT GAC CAT TCC ACA GAA AAA GGC TTC	453
Val Ser Glu Leu Ala Arg Gly Cys Asp His Ser Thr Glu Lys Gly Phe	20 25 30
	35 40
CGC TAT GTT ACT ATC TTT AGG TGT AAC CAC TTA GCG TTA AAG CCA CTT	501
Arg Tyr Val Thr Ile Phe Arg Cys Asn His Leu Ala Leu Lys Pro Leu	45 50
	35 40

203/235

FIG.29B

549	TCC GCT ATG TTA CTA TCT TTA GGT GTA ACA TCT ATT CCA CAA TCT GTT Ser Ala Met Leu Leu Ser Leu Gly Val Thr Ser Ile Pro Gln Ser Val	55 60 65
597	TTA GCA AGC GGC TTA CAA GGA ATG GAT GTA GTA CAC GGC ACA GCC ACT Leu Ala Ser Gly Leu Gln Gly Met Asp Val Val His Gly Thr Ala Thr	70 75 80
645	ATG CAA GTA GAT GGT AAT AAA ACC ATT ATC CGC AAC AGT GTT GAC GCT Met Gln Val Asp Gly Asn Lys Thr Ile Ile Arg Asn Ser Val Asp Ala	85 90 95
693	ATC ATT AAT TGG AAA CAA TTT AAC ATC GAC CAA AAT GAA ATG GTG CAG Ile Ile Asn Trp Lys Gln Phe Asn Ile Asp Gln Asn Glu Met Val Gln	100 105 110
741	TTT TTA CAA GAA AAC AAC AAC TCC GCC GTA TTC AAC CGT GTT ACA TCT Phe Leu Gln Glu Asn Asn Asn Ser Ala Val Phe Asn Arg Val Thr Ser	115 120 125 130
789	AAC CAA ATC TCC CAA TTA AAA GGG ATT TTA GAT TCT AAC GGA CAA GTC Asn Gln Ile Ser Gln Leu Lys Gly Ile Leu Asp Ser Asn Gly Gln Val	135 140 145
837	TTT TTA ATC AAC CCA AAT GGT ATC ACA ATA GGT AAA GAC GCA ATT ATT Phe Leu Ile Asn Pro Asn Gly Ile Thr Ile Gly Lys Asp Ala Ile Ile	150 155 160

204/235

FIG.29C

AAC ACT AAT GGC TTT ACG GCT TCT ACG CTA GAC ATT TCT AAC GAA AAC Asn Thr Asn Gly Phe Thr Ala Ser Thr Leu Asp Ile Ser Asn Glu Asn 165 170 175	885
ATC AAG GCG CGT AAT TTC ACC TTC GAG CAA ACC AAA GAT AAA GCG CTC Ile Lys Ala Arg Asn Phe Thr Phe Glu Gln Thr Lys Asp Lys Ala Leu 180 185 190	933
GCT GAA ATT GTG AAT CAC GGT TTA ATT ACT GTC GGT AAA GAC GGC AGT Ala Glu Ile Val Asn His Gly Leu Ile Thr Val Gly Lys Asp Gly Ser 195 200 205 210	981
GTA AAT CTT ATT GGT GGC AAA GTG AAA AAC GAG GGT GTG ATT AGC GTA Val Asn Leu Ile Gly Gly Lys Val Lys Asn Glu Gly Val Ile Ser Val 215 220 225	1029
AAT GGT GGC AGC ATT TCT TTA CTC GCA GCG CAA AAA ATC ACC ATC AGC Asn Gly Gly Ser Ile Ser Leu Leu Ala Gly Gln Lys Ile Thr Ile Ser 230 235 240	1077
GAT ATA ATA AAC CCA ACC ATT ACT TAC AGC ATT GCC GCG CCT GAA AAT Asp Ile Ile Asn Pro Thr Thr Tyr Ser Ile Ala Ala Pro Glu Asn 245 250 255	1125
GAA GCG GTC AAT CTG GGC GAT ATT TTT GCC AAA GCG GGT AAC ATT AAT Glu Ala Val Asn Leu Gly Asp Ile Phe Ala Lys Gly Gly Asn Ile Asn 260 265 270	1173

09/806709

205/235

FIG.29D

GTC GGT GCT GCC ACT ATT CGA AAC CAA GGT AAA CTT TCT GCT GAT TCT Val Arg Ala Ala Thr Ile Arg Asn Gln Gly Lys Leu Ser Ala Asp Ser 275 280 285 290 295	1221
GTA AGC AAA GAT AAA AGC GGC AAT ATT GTT CTT TCC GCC AAA GAG GGT Val Ser Lys Asp Lys Ser Gly Asn Ile Val Leu Ser Ala Lys Glu Gly 295 300 305	1269
GAA GCG GAA ATT GCG GGT GTA ATT TCC GCT CAA AAT CAG CAA GCT AAA Glu Ala Glu Ile Gly Gly Val Ile Ser Ala Gln Asn Gln Ala Lys 310 315 320	1317
GCG GCG AAG CTG ATG ATT ACA GCG GAT AAA GTC ACA TTA AAA ACA GGT Gly Gly Lys Leu Met Ile Thr Gly Asp Lys Val Thr Leu Lys Thr Gly 325 330 335	1365
GCA GTT ATC GAC CTT TCA GGT AAA GAA GCG GGA GAA ACT TAC CTT GGC Ala Val Ile Asp Leu Ser Gly Lys Glu Gly Gly Thr Tyr Leu Gly 340 345 350	1413
GGT GAC GAG CGC GCG GAA GGT AAA AAC GCG ATT CAA TTA GCA AAG AAA Gly Asp Glu Arg Gly Glu Gly Lys Asn Gly Ile Gln Leu Ala Lys Lys 355 360 365 370	1461
ACC TCT TTA GAA AAA GCG TCA ACC ATC AAT GTA TCA GCG AAA GAA AAA Thr Ser Leu Glu Lys Gly Ser Thr Ile Asn Val Ser Gly Lys Glu Lys 375 380 385	1509

206/235

FIG.29E

GCC GGA CGC GCT ATT GIG TGG GGC GAT ATT GCG TTA ATT GAC GGC AAT Gly Gly Arg Ala Ile Val Trp Gly Asp Ile Ala Leu Ile Asp Gly Asn 390 395 400	1557
ATT AAC GCT CAA GGT AGT GGT GAT ATC GCT AAA ACC GGT GGT TTT GTG Ile Asn Ala Gln Gly Ser Gly Asp Ile Ala Lys Thr Gly Gly Phe Val 405 410 415	1605
GAG ACA TCG GCG CAT TAT TTA TCC ATT GAC AGC AAT GCA ATT GTT AAA Glu Thr Ser Gly His Tyr Leu Ser Ile Asp Ser Asn Ala Ile Val Lys 420 425 430	1653
ACA AAA GAG TGG TTG CTA GAC CCT GAT GAT GTA ACA ATT GAA GCC GAA Thr Lys Glu Trp Leu Leu Asp Pro Asp Asp Val Thr Ile Glu Ala Glu 435 440 445 450	1701
GAC CCC CTT CCG AAT AAT ACC GGT ATA AAT GAT GAA TTC OCA ACA GGC Asp Pro Leu Arg Asn Asn Thr Gly Ile Asn Asp Glu Phe Pro Thr Gly 455 460 465	1749
ACC GGT GAA GCA AGC GAC CCT AAA AAT AGC GAA CTC AAA ACA ACG Thr Gly Glu Ala Ser Asp Pro Lys Lys Asn Ser Glu Leu Lys Thr Thr 470 475 480	1797
CTA ACC AAT ACA ACT ATT TCA AAT TAT CTG AAA AAC GCC TGG ACA ATG Leu Thr Asn Thr Thr Ile Ser Asn Tyr Leu Lys Asn Ala Trp Thr Met 485 490 495	1845

207/235

FIG.29F

AAT ATA ACG GCA TCA AGA AAA CTT ACC GGT AAT AGC TCA ATC AAC ATC Asn Ile Thr Ala Ser Arg Lys Leu Thr Val Asn Ser Ser Ile Asn Ile 500 505 510	1893
GGA AGC AAC TCC CAC TTA ATT CTC CAT AGT AAA GGT CAG CGT GGC GGA Gly Ser Asn Ser His Leu Ile Leu His Ser Lys Gly Gln Arg Gly Gly 515 520 525	1941
GGC GTT CAG ATT GAT GGA GAT ATT ACT TCT AAA GGC GGA AAT TTA ACC Gly Val Gln Ile Asp Gly Asp Ile Thr Ser Lys Gly Gly Asn Leu Thr 535 540 545	1989
ATT TAT TCT GGC GGA TGG GTT GAT GTT CAT AAA AAT ATT ACG CTT GAT Ile Tyr Ser Gly Gly Trp Val Asp Val His Lys Asn Ile Thr Leu Asp 550 555 560	2037
CAG GGT TTT TTA AAT ATT ACC GCC GCT TCC GTA GCT TTT GAA GGT GGA Gln Gly Phe Leu Asn Ile Thr Ala Ala Ser Val Ala Phe Glu Gly Gly 565 570 575	2085
AAT AAC AAA GCA CGC GAC GCG GCA AAT GCT AAA ATT GTC GCC CAG GGC Asn Asn Lys Ala Arg Asp Ala Ala Asn Ala Lys Ile Val Ala Gln Gly 580 585 590	2133
ACT GTA ACC ATT ACA GGA GAG GGA AAA GAT TTC AGG GCT AAC AAC GTA Thr Val Thr Ile Thr Gly Glu Gly Lys Asp Phe Arg Ala Asn Asn Val 595 600 605 610	2181

208/235

FIG.29G

TCT TTA AAC GGA ACG GGT AAA GGT CTG AAT ATC ATT TCA TCA GTG AAT Ser Leu Asn Gly Thr Gly Lys Gly Leu Asn Ile Ile Ser Ser Val Asn 615 620 625	2229
AAT TTA ACC CAC AAT CTT AGT GGC ACA ATT AAC ATA TCT GGG AAT ATA Asn Leu Thr His Asn Leu Ser Gly Thr Ile Asn Ile Ser Gly Asn Ile 630 635 640	2277
ACA ATT AAC CAA ACT ACG AGA AAG AAC ACC TCG TAT TCG CAA ACC AGC Thr Ile Asn Gln Thr Thr Arg Lys Asn Thr Ser Tyr Trp Gln Thr Ser 645 650 655	2325
CAT GAT TCG CAC TGG AAC GTC AGT GCT CTT AAT CTA GAG ACA GGC GCA His Asp Ser His Trp Asn Val Ser Ala Leu Asn Leu Glu Thr Gly Ala 660 665 670	2373
AAT TTT ACC TTT ATT AAA TAC ATT TCA ACC AAT AGC AAA GGC TTA ACA Asn Phe Thr Phe Ile Lys Tyr Ile Ser Ser Ser Asn Ser Lys Gly Leu Thr 675 680 685 690	2421
ACA CAG TAT AGA AGC TCT GCA GGG GTC AAT TTT AAC GGC GTA AAT GGC Thr Gln Tyr Arg Ser Ser Ala Gly Val Asn Phe Asn Gly Val Asn Gly 695 700 705	2469
AAC ATG TCA TTC AAT CTC AAA GAA GGA GCG AAA GTT AAT TTC AAA TTA Asn Met Ser Phe Asn Leu Lys Glu Gly Ala Lys Val Asn Phe Lys Leu 710 715 720	2517

209/235

FIG.29H

AAA CCA AAC GAG AAC ATG AAC ACA AGC AAA CCT TTA CCA ATT CGG TTT Lys Pro Asn Glu Asn Met Asn Thr Ser Lys Pro Leu Pro Ile Arg Phe 725 730 735	2565
TTA GCC AAT ATC ACA GCC ACT GGT GGT GGC TCT GTT TTT TTT GAT ATA Leu Ala Asn Ile Thr Ala Thr Gly Gly Ser Val Phe Phe Asp Ile 740 745 750	2613
TAT GCC AAC CAT TCT GGC AGA GGG GCT GAG TTA AAA ATG AGT GAA ATT Tyr Ala Asn His Ser Gly Arg Gly Ala Glu Leu Lys Met Ser Glu Ile 755 760 765 770	2661
AAT ATC TCT AAC GGC GCT AAT TTT ACC TTA AAT TCC CAT GTT CGC GGC Asn Ile Ser Asn Gly Ala Asn Phe Thr Leu Asn Ser His Val Arg Gly 775 780 785	2709
GAT GAC GCT TTT AAA ATC AAC AAA GAC TTA ACC ATA AAT GCA ACC AAT Asp Asp Ala Phe Lys Ile Asn Lys Asp Leu Thr Ile Asn Ala Thr Asn 790 795 800	2757
TCA AAT TTC AGC CTC AGA CAG ACG AAA GAT GAT TTT TAT GAC GGG TAC Ser Asn Phe Ser Leu Arg Gln Thr Lys Asp Asp Phe Tyr Asp Gly Tyr 805 810 815	2805
GCA CGC AAT GGC ATC AAT TCA ACC TAC AAC ATA TCC ATT CTG GGC GGT Ala Arg Asn Ala Ile Asn Ser Thr Tyr Asn Ile Ser Ile Leu Gly Gly 820 825 830	2853

210/235

FIG.29I

AAT GTC ACC CTT GGT GGA CAA AAC TCA AGC AGC AGC ATT ACG GCG AAT Asn Val Thr Leu Gly Gly Gln Asn Ser Ser Ser Ile Thr Gly Asn 835 840 845 850	2901
ATT ACT ATC GAG AAA GCA AAT GGT ACG CTA GAA GCC AAT AAC GCC Ile Thr Ile Glu Lys Ala Ala Asn Val Thr Leu Glu Ala Asn Asn Ala 855 860 865	2949
CCT AAT CAG CAA AAC ATA AGG GAT AGA GTT ATA AAA CTT GGC AGC TTG Pro Asn Gln Gln Asn Ile Arg Asp Arg Val Ile Lys Leu Gly Ser Leu 870 875 880	2997
CTC GTT AAT GGG AGT TTA AGT TTA ACT GGC GAA AAT GCA GAT ATT AAA Leu Val Asn Gly Ser Leu Ser Leu Thr Gly Glu Asn Ala Asp Ile Lys 885 890 895	3045
GGC AAT CTC ACT ATT TCA GAA AGC GCC ACT TTT AAA GGA AAG ACT AGA Gly Asn Leu Thr Ile Ser Glu Ser Ala Thr Phe Lys Gly Lys Thr Arg 900 905 910	3093
GAT ACC CTA AAT ATC ACC GGC AAT TTT ACC AAT AAT GGC ACT GCC GAA Asp Thr Leu Asn Ile Thr Thr Gly Asn Phe Thr Asn Asn Gly Thr Ala Glu 915 920 925 930	3141
ATT AAT ATA ACA CAA GGA GTG GTA AAA CTT GGC AAT GTT ACC AAT GAT Ile Asn Ile Thr Gln Gly Val Val Lys Leu Gly Asn Val Thr Asn Asp 935 940 945	3189

211/235

FIG.29J

GGT GAT TTA AAC ATT ACC ACT CAC GCT AAA CGC AAC CAA AGA AGC ATC Gly Asp Leu Asn Ile Thr Thr His Ala Lys Arg Asn Gln Arg Ser Ile 950 955 960	3237
ATC GGC GGA GAT ATA ATC AAC AAA GGA AGC TTA AAT ATT ACA GAC Ile Gly Gly Asp Ile Ile Asn Lys Lys Gly Ser Leu Asn Ile Thr Asp 965 970 975	3285
AGT AAT AAT GAT GCT GAA ATC CAA ATT GGC GGC AAT ATC TCG CAA AAA Ser Asn Asn Asp Ala Glu Ile Gln Ile Gly Gly Asn Ile Ser Gln Lys 980 985 990	3333
GAA GGC AAC CTC ACG ATT TCT TCC GAT AAA ATT AAT ATC ACC AAA CAG Glu Gly Asn Leu Thr Ile Ser Ser Asp Lys Ile Asn Ile Thr Lys Gln 995 1000 1005 1010	3381
ATA ACA ATC AAA AAG GGT ATT GAT GGA GAG GAC TCT AGT TCA GAT GCG Ile Thr Ile Lys Lys Gly Ile Asp Gly Glu Asp Ser Ser Asp Ala 1015 1020 1025	3429
ACA AGT AAT GCC AAC CTA ACT ATT AAA ACC AAA GAA TTG AAA TTG ACA Thr Ser Asn Ala Asn Leu Thr Ile Lys Thr Lys Glu Leu Lys Leu Thr 1030 1035 1040	3477
GAA GAC CTA AGT ATT TCA GGT TTC AAT AAA GCA GAG ATT ACA GCC AAA Glu Asp Leu Ser Ile Ser Gly Phe Asn Lys Ala Glu Ile Thr Ala Lys 1045 1050 1055	3525

09/806709

212 / 235

FIG.29K

GAT GGT AGA GAT TTA ACT ATT GGC AAC AGT AAT GAC GGT AAC AGC GGT Asp Gly Arg Asp Leu Thr 1065 1060	3573
GCC GAA GCC AAA ACA GTA ACT TTT AAC AAT GTT AAA GAT TCA AAA ATC Ala Glu Ala Lys Thr Val Thr Phe Asn Asn Val Lys Asp Ser Lys Ile 1075 1080 1085	3621
TCT GCT GAC GGT CAC AAT GGT ACA CTA AAT AGC AAA GTG AAA ACA TCT Ser Ala Asp Gly His Asn Val Thr Leu Asn Ser Lys Val Lys Thr Ser 1095 1100	3669
AGC AGC AAT GGC CGA CGT GAA AGC AAT AGC GAC AAC GAT ACC GGC TTA Ser Ser Asn Gly Gly Arg Glu Ser Asn Ser Asp Asn Asp Thr Gly Leu 1110 1115	3717
ACT ATT ACT GCA AAA AAT GTA GAA GTA AAC AAA GAT ATT ACT TCT CTC Thr Ile Thr Ala Lys Asn Val Glu Val Asn Lys Asp Ile Thr Ser Leu 1125 1130 1135	3765
AAA ACA GTA AAT ATC ACC GCG TCG GAA AAG GTT ACC ACC ACA GCA GGC Lys Thr Val Asn Ile Thr Ala Ser Glu Lys Val Thr Thr Ala Gly 1140 1145 1150	3813
TCG ACC ATT AAC GCA ACA AAT GGC AAA GCA AGT ATT ACA ACC AAA ACA Ser Thr Ile Asn Ala Thr Asn Gly Lys Ala Ser Ile Thr Thr Lys Thr 1155 1160 1165 1170	3861

213/235

FIG.29L

GGT GAT ATC AGC GGT ACG ATT TOC GGT AAC ACG GTA AGT GTT AGC GCG Gly Asp Ile Ser Gly Thr Ile Ser Gly Asn Thr Val Ser Val Ser Ala 1175 1180 1185	3909
ACT GGT GAT TTA ACC ACT AAA TOC GGC TCA AAA ATT GAA GCG AAA TCG Thr Gly Asp Leu Thr Thr Lys Ser Gly Ser Lys Ile Glu Ala Lys Ser 1190 1195 1200	3957
GGT GAG GCT AAT GTA ACA AGT GCA ACA GGT ACA ATT GGC GGT ACA ATT Gly Glu Ala Asn Val Thr Ser Ala Thr Gly Thr Ile Gly Gly Thr Ile 1205 1210 1215	4005
TOC GGT AAT ACG GTA AAT GTT ACG GCA AAC GCT GGC GAT TTA ACA GTT Ser Gly Asn Thr Val Asn Val Thr Ala Asn Ala Gly Asp Leu Thr Val 1220 1225 1230	4053
GGG AAT GGC GCA GAA ATT AAT GCG ACA GAA GGA GCT GCA ACC TTA ACC Gly Asn Gly Ala Glu Ile Asn Ala Thr Glu Gly Ala Ala Thr Leu Thr 1235 1240 1245 1250	4101
GCA ACA GCG AAT ACC TTG ACT ACT GAA GCC GGT TCT AGC ATC ACT TCA Ala Thr Gly Asn Thr Leu Thr Thr Glu Ala Gly Ser Ile Thr Ser 1255 1260 1265	4149
ACT AAG GGT CAG GTA GAC CTC TTG GCT CAG AAT GGT AGC ATC GCA GGA Thr Lys Gly Gln Val Asp Leu Leu Ala Gln Asn Gly Ser Ile Ala Gly 1270 1275 1280	4197

214/235

FIG.29M

AGC ATT AAT GCT GCT AAT GIG ACA TTA AAT ACT ACA GGC ACC TTA ACC Ser Ile Asn Ala Ala Val Thr Leu Asn Thr Thr Gly Thr Leu Thr 1285 1290 1295	4245
ACC GIG GCA GGC TCG GAT ATT AAA GCA ACC AGC GGC ACC TTG GTT ATT Thr Val Ala Gly Ser Asp Ile Lys Ala Thr Ser Gly Thr Leu Val Ile 1300 1305 1310	4293
AAC GCA AAA GAT GCT AAG CTA AAT GGT GAT GCA TCA GGT GAT AGT ACA Asn Ala Lys Asp Ala Lys Leu Asn Gly Asp Ala Ser Gly Asp Ser Thr 1315 1320 1325 1330	4341
GAA GIG AAT GCA GTC AAC GCA AGC GGC TCT GGT AGT GTG ACT GCG GCA Glu Val Asn Ala Val Asn Ala Ser Gly Ser Gly Ser Val Thr Ala Ala 1335 1340 1345	4389
ACC TCA AGC AGT GIG AAT ATC ACT GGG GAT TTA AAC ACA GTA AAT GCG Thr Ser Ser Ser Val Asn Ile Thr Gly Asp Leu Asn Thr Val Asn Gly 1350 1355 1360	4437
TTA AAT ATC ATT TCG AAA GAT GGT AGA AAC ACT GIG CCG TTA AGA GGC Leu Asn Ile Ile Ser Lys Asp Gly Arg Asn Thr Val Arg Leu Arg Gly 1365 1370 1375	4485
AAG GAA ATT GAG GIG AAA TAT ATC CAG CCA GGT GTA GCA AGT GTA GAA Lys Glu Ile Glu Val Lys Tyr Ile Gln Pro Gly Val Ala Ser Val Glu 1380 1385 1390	4533

215/235

FIG.29N

GAA GTA ATT GAA CCG AAA CCG GTC CTT GAA AAA GTA AAA GAT TTA TCT Glu Val Ile Glu Ala Lys Arg Val Leu Glu Lys Val Lys Asp Leu Ser 1395 1400 1405 1410	4581
GAT GAA GAA AGA ACA TTA GCT AAA CTT GGT GTA AGT GCT GTA CGT Asp Glu Glu Arg Glu Thr Leu Ala Lys Leu Gly Val Ser Ala Val Arg 1415 1420 1425	4629
TTT GTT GAG CCA AAT AAT ACA ATT ACA GTC AAT ACA CAA AAT GAA TTT Phe Val Glu Pro Asn Asn Thr Ile Thr Val Asn Thr Gln Asn Glu Phe 1430 1435 1440	4677
ACA ACC AGA CCG TCA AGT CAA GTC ATA ATT TCT GAA GGT AAG GCG TGT Thr Thr Arg Pro Ser Ser Gln Val Ile Ile Ser Glu Gly Lys Ala Cys 1445 1450 1455	4725
TTC TCA AGT GGT AAT GGC GCA CGA GTA TGT ACC AAT GTT GCT GAC GAT Phe Ser Ser Gly Asn Gly Ala Arg Val Cys Thr Asn Val Ala Asp Asp 1460 1465 1470	4773
GCA CAG CCG TAGTCAGTAA TTGACAAGGT AGATTTCATC CIGCAATGAA Gly Gln Pro 1475	4822
GTGATTTTAT TTTCGTATTA TTACTGTGT GGGTTAAAGT TCAGTACCGG CTTTACCCAT	4882
CTTGTA AAAA ATTACCGAGA ATACAATAAA GATTTTTTAA CAGGTATTA TTATG	4937

216/235

FIG.30A Alignment of HMW proteins

```

      10      20      30      40      50
MNKIYRLKFSKRLNALVAVSELARGCDHSTKGEKPARMKVRHLALKPLSAMLIS
.....

     110     120     130     140     150
NWKQFNIDQNEVMQFLQENNSAVFNRVTSNQISQLKGILDSNGQVFLINPNGITI
.....

     210     220     230     240     250
GLITVGKDGSVNLIGGKVKNEGVISVNGGSISLLAGQKITISDIINPTITYSIAAP
.....

     310     320     330     340     350
LSAKEGEAEIGGVISAQNQAKGGKLMITGDKVTLKTGAVIDLSGKEGGETYLGGD
.....

     410     420     430     440↓     450
GNINAQGSGLIAKTGGFVETSGHDLFIKDNAIVDAKEWLLDPDNVSINAETAGRSN
.....Y.S.DS....KT.....D.T.E..DPL.N.
.....T.E.PSYS.G.
.....E..Y...GD....D
.....G.SE.ND
.....ENPSTE.ND
.....E.T.G.GDV...D
.....D...D.PS.E.TD
.....N.VKG.ELQND
.....D.T.A.GAP..ND
.....PAL..TE
.....DIN.VNGSNIDAQ

     510     520     530     540     550
ANQRIYVNSSINL-SNGSLTLWSEGRSGG-GVEINNDITTGDDTRGANLTIYSGGW
.SRKLT.....IG..SH.I.H.K.QR...-..Q.DG...S----K.G.....
.KKNLT.....IGDSSH.I.H...KNN.-..K.KE...S----N.G....Q....
..NE.R...D..IGG.SH.....SKNKNS-..L..GN..S---.ANG.....S..
.RK..T...D..IKDSSH.I....NDNSS-..D.KGN..S---.T.GS.....S..
.TDN.....IGDS.H.I.SGG..N...-..K..KN..S---T.GS...N.K..
.KNK.L...D..IKE.SH.I.W..RDGNS-..Q.DGN..S---AT.GS..V..S..

```

217/235

FIG.30B

60	70	80	90	100	
LGVTSIPOQSVLASGLQGM DVVHGTATMQVDGNKTIIRNSVDAII					12-1
.....					12-2
160	170	180	190	200	
GKDAIINTNGFTASTLDISNENIKARNFTFEQTKDKALAEIVNH					12-1
.....					12-2
260	270	280	290	300	
ENEAVNLGDIFAKGGINIVRAATIRNQGKLSADSVSKDKSGNIV					12-1
.....					12-2
360	370	380	390	400	
ERGEKNGIQLAKKTSLEKGSTINVSQKEKGGRATVWGDIALID					12-1
.....					12-2
460	470	480	490	500	
TS-EDDEY-TGSCNSASTPKRNKE--KTTLNTTLESILKKGTFV-NIT					12-1
.G-IN..FP..TGE-..D..K.S.L-.....ISNY..NAWIM-...					12-2
AG-I.S.FPG...TKE.-..T.G.Q-P.V...E.ISNY..S..W.M...					15-1
.NL.NE...-..T.E..D.....NNT-.....S...K..AR.S.-...					15-2
A.-PTEDFP..A.GKDN-..K.AHN-.P..I.....R..SGN...-...					Joyc-1
SN-..L...-..T.ENINN..V.NQS-.K...SSI..N.....S.-...					Joyc-2
D.-S.TAFP..T.ERN.-..T.AQN-RP.I...S..Q...N....-...					LCDC2-1
.G-..V...-..T.ADINHQ.Q.S.T-.S.....GM..R.L.-...					LCDC2-2
LV-V-----R.D--..IE.K.APT-...IHAGSI.QS.M..GA.-..S					PMH1-1
G.-V..FFP..R.DD..NA.T.HPD-.P.....V.NA..NN...-...					PMH1-2
STPMNN..-D.P.QINY--K..PS-LS.....R...RN.S.-...					K1-1
LQP-----R.D---..NKVSAEGL.SIN.A..STA.Q..IE.-..S					K21-1
560	570	580	590		
VDVHKNISLGAQGNINITAKQ-DIAFEKGS-N-----QVITGQ					12-1
.....T.D-..FL....AS--V...G.N-.KARDAANAK.VA.					12-2
.....T..T-.TL....G-S.T..GNGTEKARNASSAQ..A.					15-1
..I....T.ES-.RL...T.EG.V....N------LT....					15-2
I.I....T.NS-.LL...T..G.....N------PT....					Joyc-1
..I.S.....T-.FL...SNG-SV....ADKDKARSAADAQ.VA.					Joyc-2

218 / 235

FIG.30C

.RNK.R...T..IGDS.H...YKKRKNRSD.IQ..K...S---.G.S-...N.DD.
 .TNKVN.TTD..V-Y..A...H..RD-----GN..S---EKNG....KA.S.
 .KNK.T...D..IKGGAH...Y.KNNKKS-S.K..GN..S---.TNG.....S..
 .TKT.T...D..IGDSSH.....QGR.-..NVTGN..S---.TNG.....
 .TKNVT..ADVDV-K..T.V.H.QRN-----K..GN..S---.QNG....KT..K

600 610 620 630
 G--TITSGNQ-KGFRFNNVSLNGTSGSLQFTTKRTN-----KYAI'INKFEGTLNI
 .TV...GEG--.D..A.....K..NISSVN.-----L.HNLS..I..
 .--...NIGDQ.QL.L....I....I..N.VSIQP.-----TSHR.D.E.I..
 .--...A..N-.....E.....T..L.NLS.PQ---KNNSLV..Y.N.....
 .--...A..G-.....E.A....I.T..L.NI..DL---GNNFQ.I.F.N.....
 .--I.NLTGEN.T..L.....V.Q..SI.SNVG.-----Q.H..D.EI..
 .--...ASK--.....D..T.S.VKK.FL.KYSQ.---NNKDSNFE.H.R.....
 .--...LTGEN.T..L.....N..SIISTAS.-----LSHRLD.EI.V
 .--N...NKDG.QL.L.....A..N.IANQN.-----F.HNIS.AI..
 .--V...VG-E.....V.A..R.VGQKNISSNSWRENT.K.R.D.N...
 .--..AVN.K-.....D..T....G..S.KYIE.G--NRDSNFE.H-.R.R...
 .--N.I.NQEN.QL..S.....M.A..T..ANKG.-----H.H..D.....

690 700 710 720 730
 DSRGS-----DSAGTLTQPYNLNGISFNKDTTFNVERNARVNFEDIKAPIG-INKYS
 S.---NSKGLTTQYRSSA.V.F..VNG.M-S-.LKEG.K...KL.-.NE-NMNT.
 LNNNH-GRETSR-YRKGCGVIFRSPTGHTN--.T.KQGSVA..SF..KND-T.HAN
 .-----SAE.GSAP.LSS.T.....TT.....NK..K...N.....T..Q.N
D-----T....NT.....I.D.KQ.GA.T.....-V.NNR
 E.NRFGPTTPLRS---SGGVFF..TNG.MVL--.GT.S..L.NL.-.NE-NT.N.
 .NS.SRP--SPG..P.YRRSG.....N..V...ASGSA...S..P..V-S.VHD
 N.ARNGDVR----GRSFAGVIF.AKGLTTS--...KKGST.D.KL.-.NSGYNSQK
 .NHS---TNSSDSRSFAGVKFH.KNNEMK--..IGN..KAE.RL.-.NEKTTPNR
 .AS.ISSGNQDDITNRG-----T..GEN...IAQGSTA..H..TSVMTP.PN.
 ..SGSAS--SPG..P.NAQSG.....N..V..IAASSA...N..P..V-DKVTN

 ..NRSVA--LNSGSRSFAGVKFY.KNNEMK--..IGD..N.E.KL.SNDNTSNNKP

790 800 810 820 830
 RFKTSGSTKTGFSIEKDLTLNATGGNITLLQVEGT--DGMIGK-GIVAKKNITFEG
 TLNSHVRGDDA.K.N....I...NS.FS.R.TKDDFY..YARN-A.NSTY..SIL.

219/235

FIG.30D

.....T.NS-.YL...T.SG.V...Q.N-----DLT....	LCDC2-1
..I.G..T..E-.FL...SSD-SV...G.NG.KGRSSASAO.IA.	LCDC2-2
.....T..E-.FL...SG--.....N-.-----LT..A.	PMH1-1
..I....T.NT-.YL...GG-SV...AGNEKGRQVSES..KA.	PMH1-2
.....T.KS-.YL...T..G.....DKPGLS-----NLT..AK	K1-1
.....T..M-.FL...SDN-N.T....D-.-----LT..A.	K21-1
640 650 660 670 680	
SGKVNISMVLPKNESG-YDKFKGRTYWNLTSLNVSESGEFNLT--I	12-1
..NIT.NQTTR..T.--.WQTSKDSH..VSA..LETGAN.TFIKY.	12-2
..R.HVNQTT...L.--FW.VSDES...VSH.T.K.KSA.SF.KFA	15-1
..S.....IP.NAT.NW.SRY.....I.H..A..DSN....--.	15-2
.....I..-KWD-.S..R.....V.H.....GSK....--.	Joyc-1
T.N.T.NQTA.ATTA--.WN.SYDS...VST...QKNSS.TFIKRT	Joyc-2
....D.L.QARQENWN-RRHS-..SH..V.R....TNSYL.I.--.	LCDC2-1
..N.T.NQTTQQ.IE--.W.ASSDS...V..F.LR.DSK.TFIKYV	LCDC2-2
..V.T.NQTTK..AK--AWNTSYDS...VST.TL.NDAK.TFIKYV	PMH1-1
.....V..DVSGTKWH-TRIN-.....V.T...ASGSS...S--.	PMH1-2
....D.L.QARQENWN-RRHW-..SH..V.R.....NSY..V.--.	K1-1
....V.NQTT.H.-IA-PWNASADS...V.T.TLGNAQ.TFIKFV	K21-1
740 750 760 770 780	
SLNYASFNGNISVSGGGSVDFTLLASSSNVQTPGVVINSKYFNVSTGSSL	12-1
KPLPIR.LA..TAT....F.DIY.NH.---GR.AELKMSEI.I.N.ANF	12-2
-QLPIQ..S....D...K.L.CITSNY.---GRS.G.QMSSI...D..N.	15-1
N....L.....N.T.R.N....Q.....I....HL.A.K....	15-2
N.....N.T.K.....TA....F....H..A.G....	Joyc-1
KP.PLQ..A..TAI.....S.DIH.NH.---GR.AELKMNTI.I.D.T..	Joyc-2
G-.HTL...V..L...D.N.HFN.....HW.H....K.QN..A.E....	LCDC2-1
RIPIQ.QS.....R.NINT..NLT---GG..E.R.SSI...D..T.	LCDC2-2
PLPIQ.LS....T....F.DIY.NLW---GK.TELKMDSI...S..N.	PMH1-1
-...L.....L...T.N.E.N....THT.S.AI...QN....G..K.	PMH1-2
G-.HTL.K.....L...	K1-1
..D.N.HFN.....Y..Y..I.E.QNFSA.G....	K1-1
-LPIQ.LS...AT.N.T.S.DIH.NL.---ARSTEL.MSLI.I.N.VNF	K21-1
840 850 860 870 880	
GNITFGSRKAVTEIEGNVTINNANVTILIGSDFDNHQ--KPLTIKK-DVI	12-1

220/235

FIG.30E

T.NS.IRGQEA.N.S....I....SFFE.G.YSD.FNGNGFNHDA.KSTH..SIL.
 ..E.T....V..L.NN.....S.....I--.....E-.V.....T.
 E.RAE....V..L.NN.....S.....I--.....-V.....A.
 TLQSHVRKDSA.I.S....I....S.F..E.SPDSFT.KYP.R-A.SST....IS.
 ...SE...R.A.T..S.....S.N..A.I--..NLQ.-SL..N.....
 SMTAQARDRNA.E.T...VI..SNS.LSII.QNDGFDNNQKAN-A.NS.Y.V.IQ.
 TLNSHVRKYNA.E.N....I...NS.FN.R.TSDSFRN.YRNN-A.NSTH..SIL.
 NL.A....N.A.L.KNN.....EIK.....--.SR.Q.-.V..EQ..I...
 K..SE...HAA.T.KN..I.....S.N..A.I--.SNLK.-SLI.N.....
 SINSHVRGNNA.E.K...II....S.FN.K.TKDKFDNSYEKN-A.FSTH.L.IL.

890	900	910	920	930
-----	-----	-----	-----	-----

INSGNLTAGGNIVNIAGNLTVESNANFKAITNFTFNVGGLFDNKGNSNISIAGGA
 LVN.S.SLT.ENAD.K....ISES.T..GK.RD.L.IT.N.T.N.TAE.N.TQ.VV
 .E-...SLI.ASA..N...S.KE..K..GE.QDNL.IT.T.I.N.D.K.N.SQ.VV
 NR-.....VI..G.....NG..L.....N.....R...
VI..N....NNG..L.....R...
 V-E...LT.SVAD.K...SILND.T..GE.SENL.IT.N.T.N.TAD.N.KQ.VV
 -K....VT.SAI..EK....GS.K.L.NP.YS...S....Q.K.....
 VE-.E.RLV.ASA..NN..S.K.G.K...E..DNL.IT.T.T.N.T.I.DVK..A.
 VE-...SLI.ENA..N...SI.KE.I..GK.KDSL.IT.N.T.N.TAE.N.SQ.VV
 TAN.....D.DTIK.K...D.AQG.K.NGS.KNNL.IT.T.T.N.T.I.D.TQ.VV
 NK-...VT.SAI..EK....GS.K.L.NP.YS...S....Q.K.....
 SVG...NII.SNAH.D...SIAES.K.QGK..NNL.IT.T.T.N.TAD.N.KQ.VV

990	1000	1010	1020	1030
-----	------	------	------	------

TEMQIGGDVSQKEGNLTISSDKINITKQITIKAGVDGENSDSDATNNANLTIKTKE
 A.I...NI.....K.I...D.S....S.....
 A.I...NI.....N.....
 A.I...NI.....K..N.....STKSQ.....
 A.I...NI.....N....K..NK.D...STA.....
 A.I...NI..K.....K.....EGG...SPAS.....T
 A.I...NI.....N....K..NK.D...STA.....
 ..I...NI.....R.E...T.QG...GVAS.....
 A.IE...NI.....V.....S.S.STASD.....
 A.IE...NI.....V.....S.S.STASD.....
 ..I...NI.....V...ER.....N.D...NEATS.....
 A.I...NI.....R.E...DT.QG...GVAS.....

221/235

FIG.30F

..V.L.GQNSSSS.T..I..EKA.....EANNAP.Q.NIRDRV..LG-SL	12-2
..V.L.GQDSSST.T..IN.SQA.....RAYNGNGRN--.Q.--LGN.S	15-1
.....K..I...K.....E.T.A.....ND.K--...N..G--..V	15-2
.....K..I.....A.....--.....-...	Joyc-1
..VSL.GQNSSSD.K..I..KSST....KAHNSPRDFASRT.--LGNLN	Joyc-2
....LAAD.KPI..K..I.VKEG.....RSANYG.DK--SA.S.R-GN.T	LCDC2-1
..V.L.GQNSSST.T.S.N.GA.....QAHNGNDRN--.K.--FGN.S	LCDC2-2
..V.L.GQNSSSS.M..II.KRA.....EADNSH.SDNV.DR..NLGNLT	PMH1-1
....L..Q..P...K.D..VKQGT.A..RSAN....---.GAL.VNGN.-	PMH1-2
....LAAD.KPI..K..I.VKEG.....RSANYG.DK--SA.S.-RGN.T	K1-1
..V.L.GENSSSN.K..IN..SK.....QAHAGTS.LDK.ER.LTLGN.-	K21-1

940	950	960	970	980	
RFK-DIDNSKNLSITTNSSSTYRTIISGNITNKNGDLNITNEGSD					12-1
KLG-NVT.DGD.N...HAKRNQ.S..G.D.I..K.S....DSNN.					12-2
KLG-NVT.DGD.N...HAKHNQ.S..G.D.I..K.S....DSNKN					15-1
K..-..N.TSS.N....DT.....E.....A.....IDNKGN					15-2
K..-..N.TSS.N....DT.....E.....A.....IDNKGN					Joyc-1
NIQGN.T.KGG.N....AQNNQK...N.....EG.....KDSNNN					Joyc-2
H..-..N.T.S.N....D.A.....E.....S.....DNKNN					LCDC2-1
KLG-N.T.DG..N....AKNGQKSV.N.....NK.A.....N.N.					LCDC2-2
SLG-..T.DGK.N...HAK.GOKS.. <u>R</u> . <u>D</u> .I..Q.N....DNN.N					PMH1-1
NLG-NVT.DGK.N...HAK.GOKS.. <u>R</u> . <u>D</u> .I..Q.N....DNN.N					PMH1-2
I..-..E.TGS.N...K.D.NHH...K.....RK.....N.DN					K1-1
KLQG..T.NG..N....A.VNQK...N.....K.....KDIKAN					K21-1

1040	1050	1060	1070	1080	
LKLTQDLNISGFNKAETAKDGSDLTIGNIN-SADGTINAKKVTFN					12-1
....E..S.....R.....S.DGNS.AE..T....					12-2
.....-...S.....					15-1
.....K.V...S.N....SDD.GN-.S..T....					15-2
.Q..G.....D.....E.A..I...SDNNNA-.....					Joyc-1
.E..G.....N.N....KASDGNA--.....D					Joyc-2
.Q..G.....D.....E.A..I...SDNNNA-.....					LCDC2-1
....EN.....D...V..ENNN.I...N.--G.NA...T....					LCDC2-2
.T..DN.....N...I..KA--.S.NS...Q...D					PMH1-1
.TF.DN.....N...I..KA--.S.NS...Q...D					PMH1-2

223/235

FIG.30H

....N.....N.N....DNSD-.GN.D.....S	K1-1
.T..DN.....N...I..KA--.S.NS...QI..D	K21-1
1140 1150 1160 1170 1180	
HKAVSISATSGEITTKTGTINATTGNVEITAQTGSILGGIESSG	12-1
L.T.N.T.SEK-V..TA.S.....N.KAS..TK..D.S.-----	12-2
N.T.N.T.SEK-L...ADA.....V..K..D.K.EVK.T--	15-1
..T.NV..AN.G.....A.....H....Q.....KP.	15-2
..T.N.T.SEN-V...A.....I.S..V..K..D.K....N..	Joyc-1
..TIN....T.NV...ES.....A..S..V..K..D.S.-----	Joyc-2
..T.N.T.SEN-V...A.....S..V..K..D.K.KV..T..	LCDC2-1
..T.N.T.SER-.D..AD.....KL..V.SD.-...K.N..	LCDC2-2
..T.N...SE.G...A.....S..V..K.-----	PMH1-1
..T.N...SE.G...A.....S..V..K.-----	PMH1-2
..T.N.T.SEN-V...A.....S..V..K..D.K....N..	K1-1
..T.N...SE.G...A.....S..V..K..D.S.T.SGKTV	K21-1
-----	12-1
-----	12-2
-----	15-1
-----	15-2
EVKSASGNVNITASGNITLNVSNITGQNVTVTANSGAITTEGSTI	Joyc-1
-----	Joyc-2
-----	LCDC2-1
-----	LCDC2-2
-----	PMH1-1
-----	PMH1-2
EVKSASGNVNITASGNITLNVSNITGQNVTVTANSGAITTEGSTI	K1-1
-----	K21-1
1220 1230 1240 1250	
-----KGTESVTTSSQSGDIGGTISGGTIVEVKAT-ESL	12-1
-----.....N..S.S..-GD.	12-2
-----	15-1
-----I.....N...K...K..N....-N..	15-2
TQTSSKIN..K.....S.....N..S.S..-G..	Joyc-1
-----.....N..N.T..D.--	Joyc-2

224/235

FIG.30I

 NATTGDANITTQTGNINGKVESSSGSVTLIATGQTLAVGNI SGDTVTITADKGKLT
 -----TDLTTVKGGAKINATEGTATLTASSGKLT

 TTQADSKIEATEGEANVTSKTSIIGGTISGGTVEVTATEGLTTQAGSTITGTESVT

1260	1270	1280	1290	1300
TTQSN	SIKAT	TG	ANVT	SATGT
IIGGT	ISGN	IVNV	TANAG	DLTV
GN	GA	EL	IN	ATEGA
..K.G...	E.KS.....			
-----I....	TLNVS----	VSGN.	
I.K.G.E...	A.V.....	D.....	T....	EDA.K.D..G..
...AG...	E.K.....		TDN..IKD..	R.K..G..

..K.G.E...	K.....	A.....	T....	EDA.K.D..G..
-----				..KG..K.....T
-----	D.S....	K..S...	ST....	RKA.T.S....
-----	D.S....	K..S...	ST....	RKA.T.S....
...AG...	E.K.....		TDN..IKD..	R.K..G..
-----				...GD.K.....

1360	1370	1380	1390	1400
LNITG	TLTT	VKGS	NINAT	SGTL
VI	NAK	DAEL	NGAAL	GNHT
VV	NAT	NANG	SGSV	IAT
.....A..	D.K.....	K...D.S.	DS.E...V..	S.....T.A
.....E.....	A.....	K....S.D.S.....	T.V
.....A.....	S...N.....	E.S.....		
.....A..	K.E.A.....	Q.D....	DR.E..V.....	
.....A...	K.....A.....	K.D.T.S.	DR.....S.....	T.A

225/235

FIG.30J

-----SAING..A.....S.....N..K.S.I-GD.	LCDC2-1
TLAGSTIN..NG.....E...EVT.K..S.T..AG..	LCDC2-2
-----	PMH1-1
-----	PMH1-2
TQTSSKIN..K.....S.....N..S.S..G-..	K1-1
TEANSAIS.ANG..A.....S.....K..S.T.SSG..	K21-1
-----	12-1
-----	12-2
-----	15-1
TSSQSGNIGGMISGGKVEVSATKDL	15-2
-----	Joyc-1
-----	Joyc-2
-----	LCDC2-1
-----	LCDC2-2
-----	PMH1-1
-----	PMH1-2
-----	K1-1
-----	K21-1
1310 1320 1330 1340 1350	
ATLTSSGKLTTEASSHITSAGQVNLQAQDGSVAGSINAANVT	12-1
....ATGNT.....G.S...T....D.L..N..I.....	12-2
V.I.ADK.....Q...S...NN..TT.T.K...I.....	15-1
....AT.....K...S...NN.....K...IG.N.....	15-2
V...ATG.T...T..D...SN..TT.T.K.S.I.....	Joyc-1
-----...Q...S...SN..TT.T.KN..I...D....	Joyc-2
....AT.....K...S...NN.....K...IG.N.....	LCDC2-1
....A.....N.....D.....I..Q.S.....	LCDC2-2
....ATGNT.....G.S...T....D.....I..Q.S.....	PMH1-1
....ATGNT.....G.S...T....D.....I..Q.S.....	PMH1-2
V...ATG.T...T..D...SN..TT.T.K.S.I.....	K1-1
....ATK.T...-----	K21-1
1410 1420 1430 1440 1450	
TSSRVNITGDLITINGLNIISKNGINTVLLKGVKIDVKYIQPGI	12-1
...S.....N.V.....D.R...R.R.KE.E.....V	12-2
...N.....S.V.....R...V...TE.E.....V	15-1

226/235

FIG.30K

.....A.....S.....N.....E.S.....
E.S.....NE.....N..K.D.K.S..R.E.....S.....T.K
E.....K.....A.....K.D.T.S..R.E.....S.....T.K
E.....K.....A.....K.D.T.S..R.E.....S.....T.K
A..K.E.A.....Q.D...S.D.....
 -----.....D.NE.....Q..T...D.S.DR.E...V..S...N.T.K

1460	1470	1480	1490	1500
ASVDEVIEAKRILEKVKDLSDEEREALAKLGVSAVRFIEPNNTITVDTQNEFATRP				
...E.....V.....T.....V.....N.....T...				
...E.....V.....T.....N.....T...				
.....A...A..IN....T...				
...Y.....A.....				
..AN.....A.....T.....V.....N.....T...				
.....A...A..IN....T...				
..E.....V.....T.....N.....T...				
...E.....V.....T.....V...A..IN....T...				
...E.....V.....T.....V...A..IN....T...				
...N.....A.....T.....A...A..IN....T...				
..AN.....A.....T.....N.....T...				

227/235

FIG.30L

.....	15-2
.....K.....E.....	Joyc-1
...N.....S.....K...V...AE.....V	Joyc-2
.....	LCDC2-1
...S.....N.....E..R...R.R.KE.E.....V	LCDC2-2
...N.....S.....E..R...R.R.KE.....V	PMH1-1
...N.....S.....E..R...R.R.KE.....V	PMH1-2
.....K.....E.....	K1-1
...S.....S.....K...V...AE.....V	K21-1

1510	1520	1530	
LSRIVISEGRACFSNSDGATVCVNIADNGR-*			12-1
S.QVI....K....SGN..R..T.V..D.QP*			12-2
S.QVT....K....SGN..A..T.V..D.QQ*			15-1
S.QVT....KV..LIGN...I.T....IE.*			15-2
.....*			Joyc-1
S.QVT...DK....SGN..A..T.VT.DRQ*			Joyc-2
S.QVT....KV..LIGN...I.T....IE.*			LCDC2-1
S.QVT....K....SGN..A..T.V..D.QQ*			LCDC2-2
S.QVI....K....SGN..A..T.V..D.QP*			PMH1-1
S.QVI....K....SGN..A..T.V..D.QP*			PMH1-2
..QVT....KV..LIGN...I.T....IE.-*			K1-1
S.QVT....K....SGN..A..T.V..D.QQ*			K21-1

228 /235

FIG.31

Oligonucleotides used to determine whether PCR amplified *hmwA* genes were *hmw1* or *hmw2*.

		<u>SEQ ID NO</u>	
5'	TCCTTTCTGCTGGCTGATGCCCCCTA	3'	5672.SL
5'	CACGATAGGTGCTCATATTCGCC	3'	5676.SL
V G V H K N			
GGTTCATGTTTCATAAAAATAT			
3'	CCAACTACAAGTATTTTATA	5'	5742.SL
G G S L T I N S			
GGCGGAAGTTTAACTATTAACTC			
3'	CCGCCTTCAAATTCATAATTGAG	5'	5743.SL

229/235

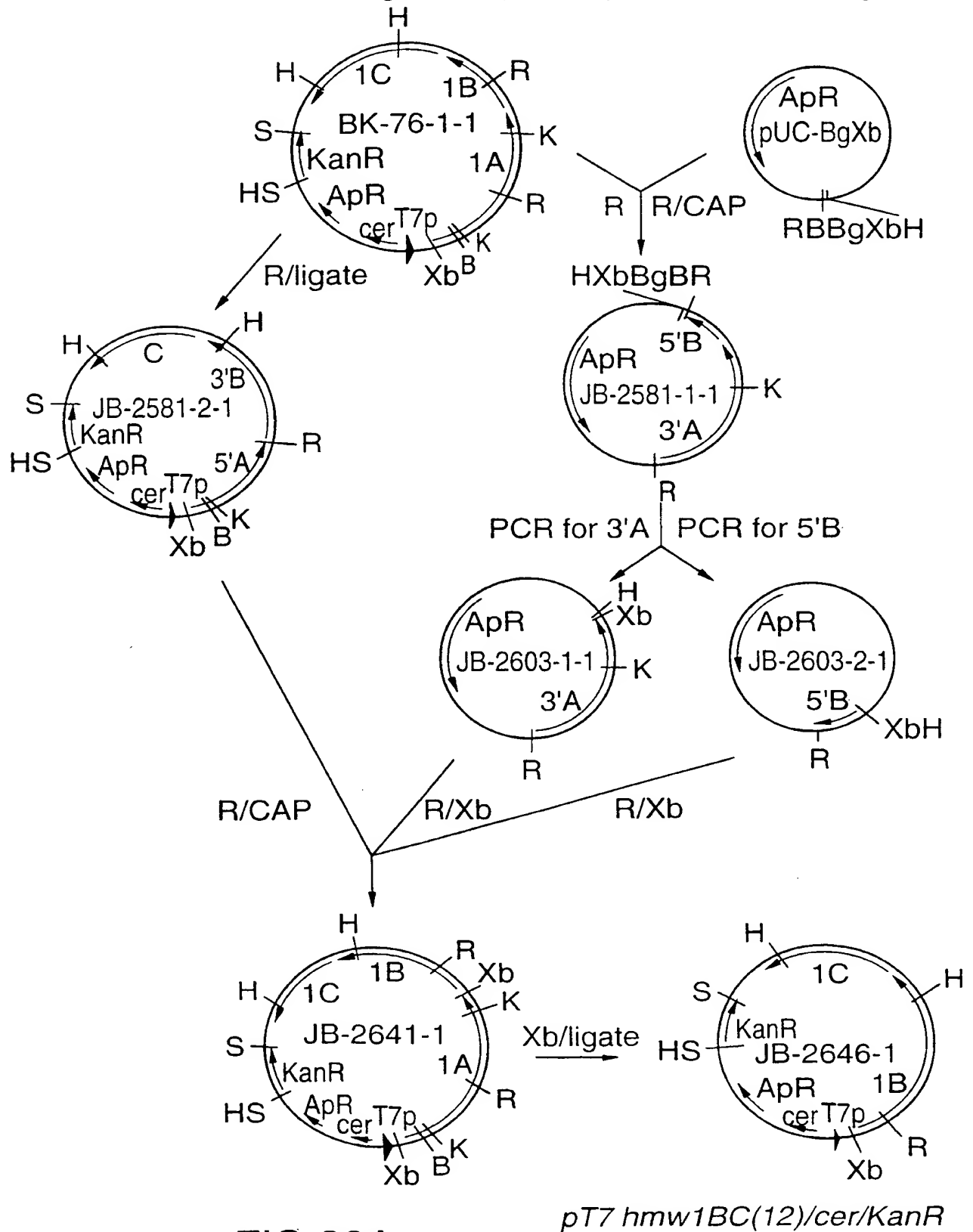
Construction of JB-2646-1, a generic expression plasmid for *hmvABC* genes

FIG.32A

SUBSTITUTE SHEET (RULE 26)

FIG.32B

Oligonucleotides used to PCR amplify the 3'-end of *hmm1A* and 5'-end of *hmm1B* to construct a generic expression vector.

3'-end of *hmm1A*

EcoR I



G V D G E N S D S D
GGTGTTCATGGGAGAAATTCGATTCAGACG

5' 3'

5947.SL

82
83

SEO ID NO

230/235

V C V N I A D N G R *

GTGTGCGTTAATATCGCTGATAACGGGGGTAG

CACACGCAATTATAGCGACTATTGGCCCGCCATCAGATCTCCGG

5'

5948.SL

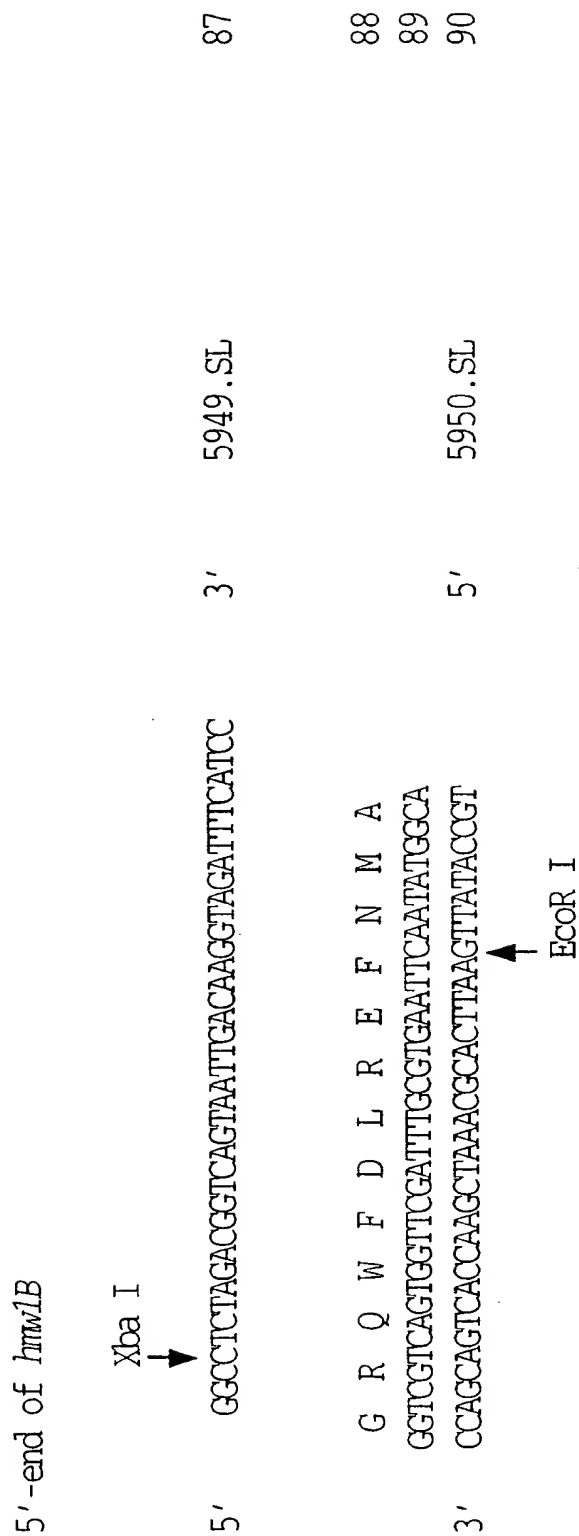
84
85
86



Xba I

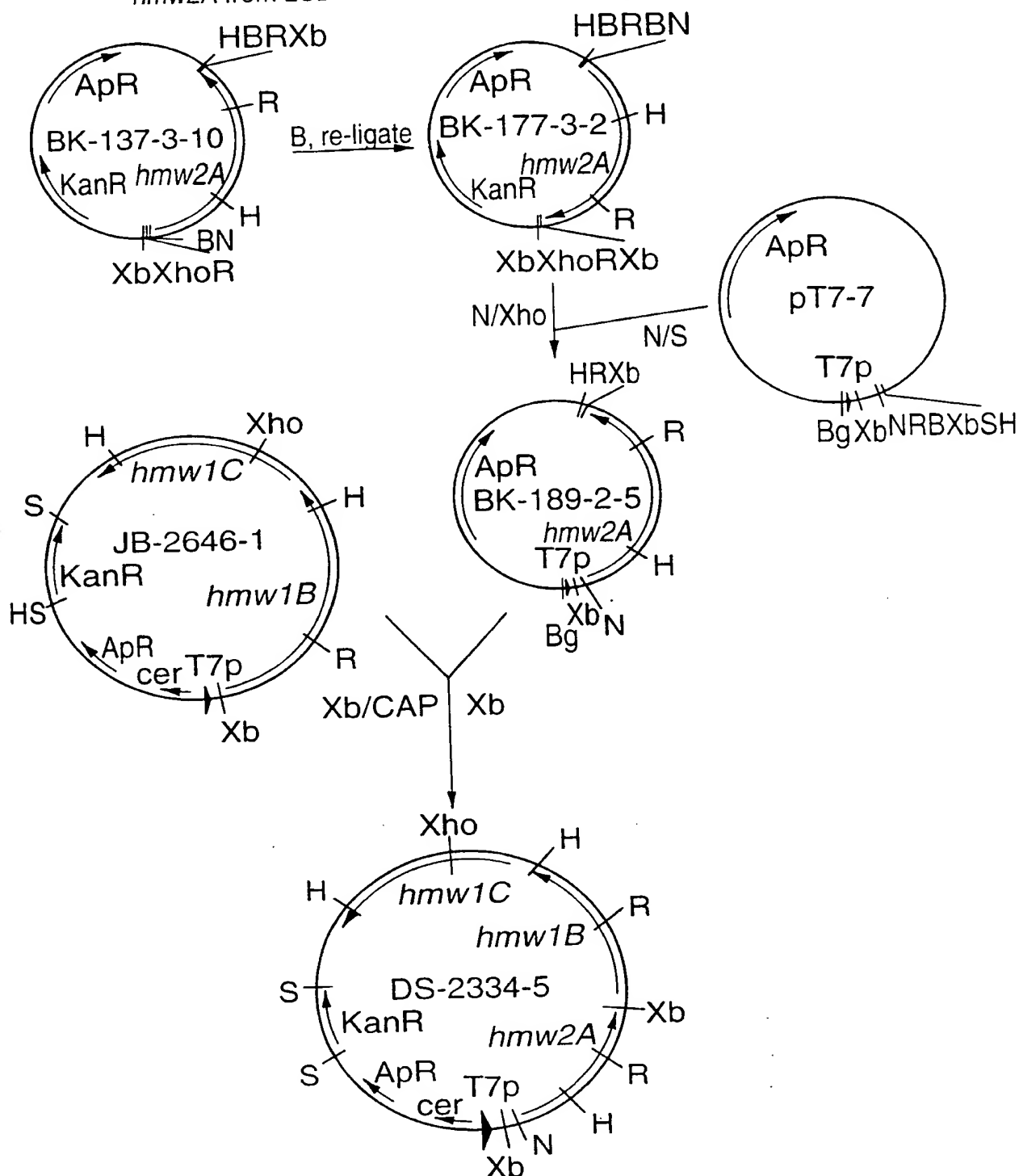
231/235

FIG.32B'



232/235

Construction of DS-2334-5, a chimeric expression plasmid with *hmw2A* from LCDC2 and *hmw1BC* from strain 12



pT7hmw2A (LCDC2)/*hmwBC* (12)/*cer*/*KanR*

FIG.33A

233/235

FIG.33B

Oligonucleotides used to PCR amplify the LCDC2 *hmm2A* gene for expression

sense

	BamH I Nde I			
	↓	↓		
		M P D D V S I D A P S A E		
5'		CGGGATCCCATATGCCGGATGATGTATCCATTCACGACCTTCGGCTGAA	3'	5972.SL

	<u>SEQ ID NO</u>
	91
	92

antisense

	A A V C T N V A D D G Q Q *		
5'	GCAGCAGTATGTACCAATGTTCCTGACGATGGACAGCAGTAGT	3'	93
3'	CGTCGTACATAGGTACACGACTGCTACCTGCTGCTCAGATCTG	5'	94
	↑		95
	XBA I		5973.SL

234/235

Construction of DS-2400-13, a pBR T7 hmwA/T7 hmwABC/cer/KanR plasmid

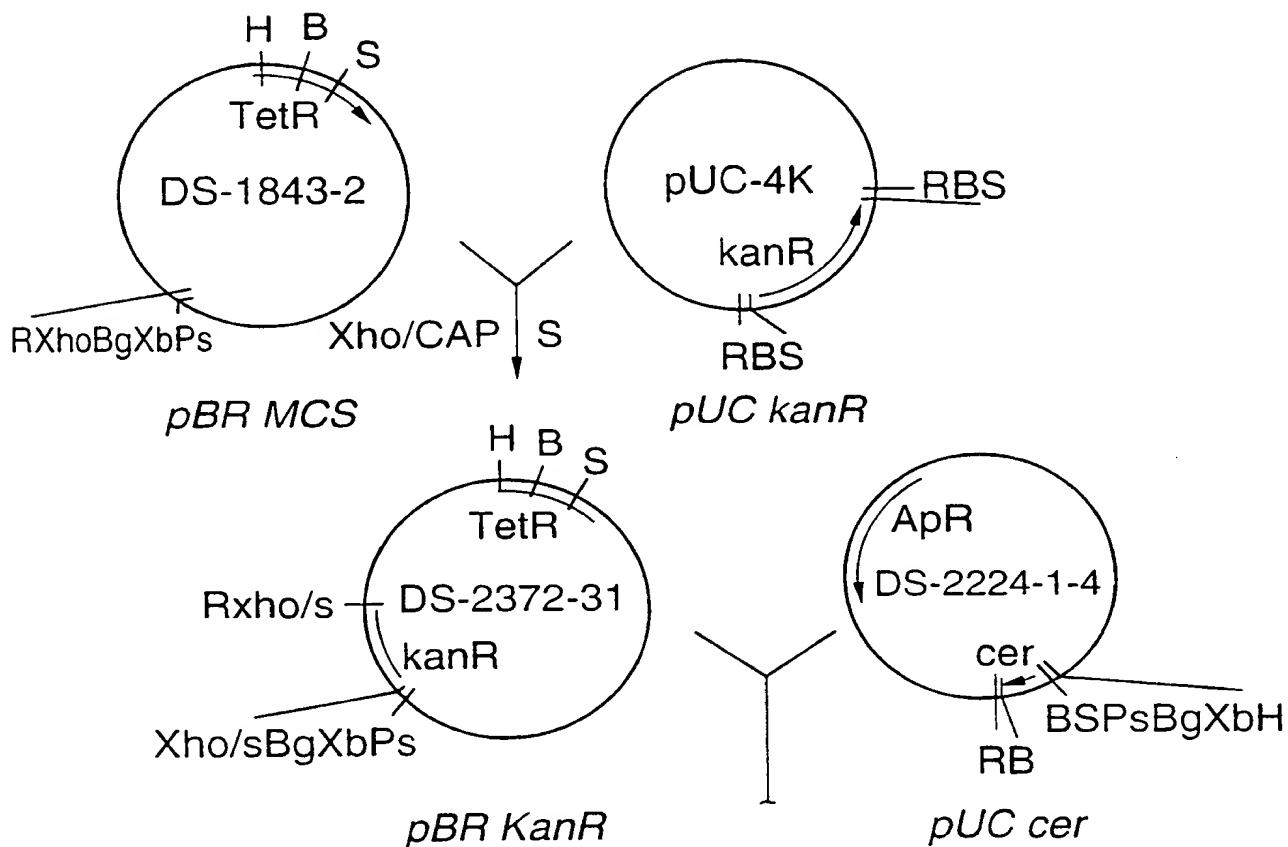


FIG.34A

235/235

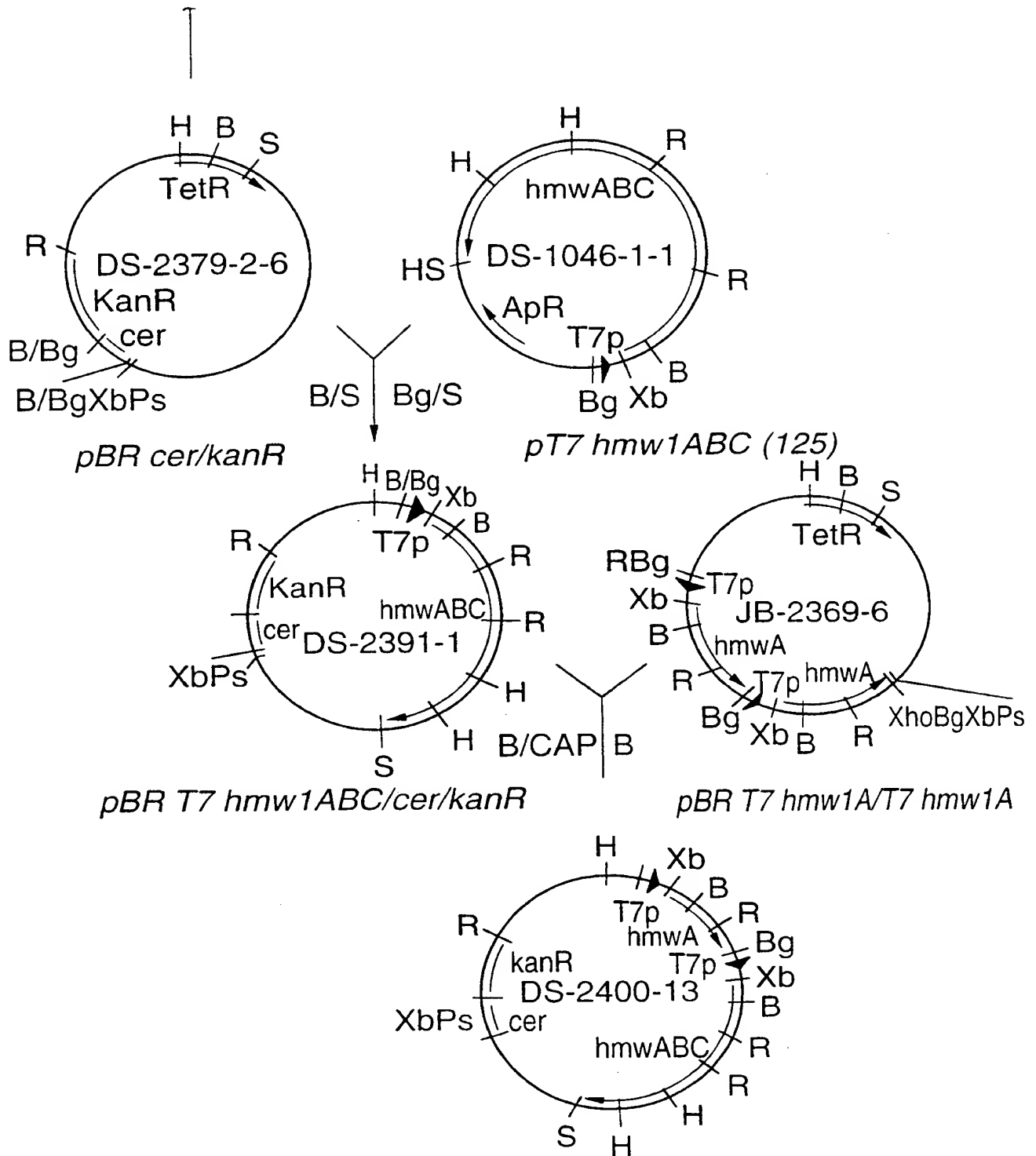


FIG.34A'

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.